



HAL
open science

Coordinated Multi-cell Beamforming for Massive MIMO: A Random Matrix Approach

Subhash Lakshminarayana, Mohamad Assaad, Merouane Debbah

► **To cite this version:**

Subhash Lakshminarayana, Mohamad Assaad, Merouane Debbah. Coordinated Multi-cell Beamforming for Massive MIMO: A Random Matrix Approach. *IEEE Transactions on Information Theory*, 2015, 61 (6), pp.3387-3412. 10.1109/TIT.2015.2421446 . hal-01237227

HAL Id: hal-01237227

<https://hal.science/hal-01237227>

Submitted on 5 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Coordinated Multi-cell Beamforming for Massive MIMO: A Random Matrix Approach

Subhash Lakshminarayana, *Member, IEEE*, Mohamad Assaad, *Member, IEEE*, and Merouane Debbah, *Senior Member, IEEE*

Abstract—We consider the problem of coordinated multi-cell downlink beamforming in massive multiple input multiple output (MIMO) systems consisting of N cells, N_t antennas per base station (BS) and K user terminals (UTs) per cell. Specifically, we formulate a multi-cell beamforming algorithm for massive MIMO systems which requires limited amount of information exchange between the BSs. The design objective is to minimize the total transmit power across all the BSs subject to satisfying the user signal to interference noise ratio (SINR) constraints. With our algorithm, the BSs need to exchange parameters which can be computed solely based on the channel statistics rather than the instantaneous CSI. We make use of tools from random matrix theory to formulate the decentralized algorithm. We also characterize a lower bound on the set of target SINR values for which the decentralized multi-cell beamforming algorithm is feasible. We further show that the performance of our algorithm asymptotically matches the performance of the centralized algorithm with full CSI sharing. Finally, we investigate the impact of imperfect CSI and pilot contamination effect on the performance of the decentralized algorithm, and propose a heuristic extension of the algorithm to accommodate these issues. Simulation results illustrate that our algorithm closely satisfies the target SINR constraints and achieves minimum power in the regime of massive MIMO systems. In addition, it also provides substantial power savings as compared to zero-forcing beamforming when the number of antennas per BS is of the same orders of magnitude as the number of UTs per cell.

Index Terms—Massive MIMO, coordinated beamforming, decentralized design, random matrix theory.

I. INTRODUCTION

Massive multiple input multiple output (MIMO) has been identified as an essential ingredient in the design of next generation cellular systems, as it provides substantial improvement in both spectral and energy efficiency [1]. It refers to the idea of scaling up the number of antennas on the base station (BS) to a few hundreds, serving many tens of user terminals (UTs) on the same resource block. The basic idea is to exploit large number of antennas to achieve greater spatial resolution and array gain, resulting in a higher throughput and greater energy efficiency.

This paper was presented in part at the IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), September, 2010. This research has been supported by the ERC Starting Grant 305123 MORE and by the French pole de competitivite SYSTEM@TIC within the project 4G in Vitro.

S. Lakshminarayana is with the Singapore University of Technology and Design (SUTD), Singapore (email: subhash@sutd.edu.sg).

M. Assaad is with the Dept. of Telecommunications, SUPELEC, France (email:mohamad.assaad@supelec.fr).

M. Debbah is with the Alcatel-Lucent Chair on Flexible Radio, SUPELEC, France (email: merouane.debbah@supelec.fr).

The idea of massive MIMO was first proposed in the seminal work of [2]. The main finding of the paper was that as the number of antennas on the BS grows without bound, the effects of fast fading and interference vanish, and the system performance is ultimately limited by only pilot contamination [3]. The other attractive feature is that simple signal processing techniques at the BS, such as the use of eigen beamforming and matched filter were optimal under this setting. Subsequently, massive MIMO systems was also studied from an energy efficiency point of view and shown to achieve dramatic improvement in this regard [4]. The impact of channel estimation, pilot contamination, and antenna correlation in massive MIMO systems was investigated in [5]. In particular, it was concluded that sophisticated beamforming techniques such as regularized zero forcing (RZF) outperform eigen beamforming in a massive MIMO setting when the number of antennas is of the same orders of magnitude as the number of users. Reference [6] addressed the question of how to select the system parameters in a massive MIMO system (number of antennas per BS, number of users, transmit power etc.) to maximize the energy efficiency. The work in [7] proposes a hierarchical interference mitigation scheme in massive MIMO systems based on a two level precoding with the objective of maximizing the system utility. Although all these studies point to impressive gains in massive MIMO both in terms of spectral efficiency and energy efficiency, constructing such large dimensional arrays can result in significant additional hardware cost of the analog front ends. Moreover, extra physical dimensions are required in order reduce the mutual coupling between the antenna elements.

Subsequently, it was proposed that in a multi-cellular environment, the gains obtained by massive MIMO systems can be replicated by using much lesser number of antennas if BSs are allowed to cooperate with each other. In this context, [8] proposed a TDD architecture based network MIMO like system with BS cooperation and zero-forcing (ZF) beamforming, and showed that massive MIMO performance can be achieved with one order of magnitude fewer antennas per active user per cell. These results motivate us to consider multi-cell cooperation in massive MIMO systems. However, since massive MIMO systems are inherently large, enabling BS cooperation in a multi-cellular environment requires tremendous amount of information exchange between them.

In order to address this issue, in this work, we propose an optimal decentralized multi-cell beamforming algorithm for massive MIMO systems that requires limited amount of information exchange between the BSs. We primarily focus

on the so called *coordinated beamforming*, in which BSs formulate their beamforming vectors taking into account the interference they cause to the neighboring cells [9]. This is accomplished by exchanging the CSI information between them. However, unlike network MIMO systems, no user data exchange takes place between the BSs. The design objective in this work is to minimize the total transmitted power subject to satisfying the user signal to noise ratio (SINR) constraints. Reference [9] provides an optimal centralized algorithm to solve this problem. However, the centralized solution demands high computational ability, and exchange of the fast fading CSI co-efficient between the BSs. Such an algorithm requires high capacity backhaul links, especially when implemented in a massive MIMO setting.

In order to overcome the heavy backhaul requirement, we propose in this work a decentralized approach to compute the multi-cell beamforming vectors. In our algorithm, the BSs must exchange parameters at the time scale of slow fading coefficients rather than the instantaneous channel realizations (fast fading coefficients). We use tools from random matrix theory (RMT) to formulate our algorithm.

The multi-cell beamforming strategy involving exchange of parameters based on channel statistics was first proposed in [10] using tools from RMT. It was shown with the help of simulations that such an algorithm performs well for large system dimensions. Similar ideas for multi-cell beamforming were subsequently proposed in [11], [12] and theoretical arguments for the asymptotic optimality of this algorithm were provided in the special case of a two cell Wyner model, with symmetric SINR constraints for all the UTs. However, the analyses in [11], [12] rely on obtaining closed form expressions for the system parameters (such as the uplink and downlink power), and are not extendable to more practical channel models. In this work, we provide a comprehensive design of the RMT based decentralized beamforming under a massive MIMO multi-cell setting with a distance based pathloss model. Further, we provide arguments for asymptotic optimality of this algorithm. Specifically, our main contributions are as follows:

- We propose a *reduced overhead* beamforming algorithm (ROBF) in a massive MIMO multi-cell setting. In our algorithm, the BSs require the knowledge of local CSI (of the UTs they are serving and also the UTs present in the other cells). In addition, the BSs have to compute parameters that depend only on the channel statistics, which they exchange between them to compute the beamforming vectors.
- Using a large system analysis, we provide closed form expression for the lower bound on the set of target SINR, for which the decentralized multi-cell beamforming algorithm is feasible.
- We prove that when the dimensions of the system become large, the achieved SINR in the uplink and downlink by the ROBF algorithm exactly match the target SINR. Moreover, we also prove that when the dimensions of the system become large, the performance of our algorithm in terms of uplink and downlink transmit powers perfectly match that of the optimal algorithm proposed in [9].
- Finally, we investigate the impact of CSI estimation and

pilot contamination on the performance of the ROBF algorithm and propose a heuristic adaptation of the ROBF algorithm that can provide better performance in the presence of pilot contamination.

In addition, our work contains several novel ideas of combining RMT results with optimization theory and uplink-downlink duality in MIMO systems, which are of independent interest.

It is worth noting that RMT results have been used extensively to assess the performance of linear beamforming strategies in multi-user MIMO systems [13], [14], [5], [8]. However, all these works consider "pre-defined" transmit strategies such as eigen beamforming, zero forcing (ZF), regularized zero forcing (RZF) etc., and use RMT as a tool for *performance analysis* of these schemes in the large dimensional regime. In contrast, in this work, we use RMT results for the purpose of *system design* (as opposed to performance analysis), i.e., design of optimal beamforming vectors in MIMO multi-cell systems¹. Such an approach is novel and is facilitated by combining RMT results with optimization theory. Moreover, optimizing the system performance imposes additional technical difficulties in applying RMT results. For e.g., implementing a power control algorithm (both in the uplink and downlink) implies that the transmit powers explicitly depend on the channels (and hence the randomness associated with the channel realization). Such dependency of transmit powers on the channel realization makes it unsuitable to apply RMT results. Our approach in this work is to first propose an algorithm that depends only on the second order statistics of the channel vectors (the path-loss in our case). Then, we apply such an algorithm to the original system set-up, and prove that this algorithm is optimal in the large system domain.

Although the theoretical results prove the optimality of ROBF algorithm in the asymptotic regime, we provide numerical results to show that the performance of ROBF algorithm closely matches that of the centralized algorithm for moderate system dimensions (when the number of antennas are comparable to the number of UTs per cell), both in terms of satisfying the user SINR targets and minimizing the downlink transmission power. Moreover, these results indicate that ROBF algorithm provides substantial transmit power savings as compared to other beamforming strategies such as zero-forcing in the regime where the number of antennas is comparable to the number of UTs.

The rest of the paper is organized as follows. We provide the system model and describe our reduced overhead multicell beamforming in section II. In section III, we provide the asymptotic analysis of the reduced overhead algorithm formulated in the previous section. In Section IV, we investigate the impact of imperfect CSI and pilot contamination on the performance of the ROBF algorithm. We summarize the simulation results in section V. Finally, we provide concluding remarks in section VI. Appendix A provides some relevant results from RMT which will be used in formulating our algorithm. Appendices B, C, D, E, F, G and H provide the proofs of some of the results stated in the paper.

¹In an unrelated context, RMT results have also been used in the context of system design in [15] and [16]

Notations: Throughout this work, we use boldface lowercase and uppercase letters to designate column vectors and matrices, respectively. For a matrix \mathbf{X} , $\mathbf{X}(p, q)$ denotes the (p, q) entry of \mathbf{X} . \mathbf{X}^T , \mathbf{X}^H , $\text{tr}(\mathbf{X})$, $\|\mathbf{X}\|$, and $\rho(\mathbf{X})$ denote the transpose, the complex conjugate transpose, the trace, the spectral norm and the spectral radius of the matrix \mathbf{X} , respectively. For two matrices \mathbf{X} and \mathbf{Y} , the notation $\mathbf{X} \leq \mathbf{Y}$ denotes element wise inequality ($\mathbf{X}(p, q) \leq \mathbf{Y}(p, q) \forall p, q$) and similarly for vectors. We denote an identity matrix of size M as \mathbf{I}_M and $\text{diag}(x_1, \dots, x_M)$ is a diagonal matrix of size M with the elements x_i on its main diagonal. We use $\mathbf{x} \sim \mathcal{CN}(\mathbf{m}, \mathbf{R})$ to state that the vector \mathbf{x} has a complex Gaussian distribution with mean \mathbf{m} and covariance matrix \mathbf{R} . We will use the notation $\xrightarrow{\text{a.s.}}$ to denote almost sure convergence. Let a_N and b_N denote a pair of infinite sequences. We write $a_N \asymp b_N$, iff $a_N - b_N \xrightarrow{\text{a.s.}} 0$. We denote the expectation of a random variable by the notation $\mathbb{E}[\cdot]$. Let $N_t, K \in \mathbb{N}^+$, we will use the notation $N_t, K \rightarrow \infty$ to denote the following condition on N_t and K , $0 < \liminf_{K \rightarrow \infty} \frac{N_t}{K} \leq \limsup_{K \rightarrow \infty} \frac{N_t}{K} < \infty$.

II. SYSTEM MODEL AND ALGORITHM DESCRIPTION

A. System Model

We consider the problem of multi-cell beamforming across N cells and K UTs per cell where each BS is equipped with N_t antennas and each UT has a single antenna. Each BS serves only the UTs in its cell. Let $\mathbf{h}_{i,j,k} \in \mathbb{C}^{N_t}$ denote the channel from the BS i to the k -th UT in cell j . We consider reciprocity between the uplink and downlink channels, and hence the TDD mode of operation, as it is the preferred mode of operation in massive MIMO systems [1]. We assume that the elements of the channel vector are independent and identically distributed (i.i.d.) with Gaussian distribution, i.e., $\mathbf{h}_{i,j,k} \sim \mathcal{CN}(0, \sigma_{i,j,k} \mathbf{I}_{N_t})$, the variance $\sigma_{i,j,k}$ of the channel depends upon the path loss model between BS i and UT(j, k). Recent works on channel measurements indicate that i.i.d. assumption is a reasonable model for massive MIMO arrays [17]. We assume that the BSs have perfect CSI of the downlink channels to all the users in the system (i.e., $\mathbf{h}_{i,n,k}, \forall n, k$). Let $\mathbf{w}_{i,j} \in \mathbb{C}^{N_t}$ denote the transmit downlink beamforming vector for the j -th UT in cell i . Likewise, let $\Lambda_{i,j}^{\text{DL}}$ denote the received SINR for the j th UT in cell i and $\gamma_{i,j}$ the corresponding target SINR. The received signal $y_{i,j} \in \mathbb{C}$ for the j th UT in cell i , is given by

$$y_{i,j} = \mathbf{h}_{i,j}^H \mathbf{w}_{i,j} x_{i,j} + \sum_{(n,k) \neq (i,j)} \mathbf{h}_{n,i,j}^H \mathbf{w}_{n,k} x_{n,k} + z_{i,j}$$

where $x_{i,j} \in \mathbb{C}$ represents the information signal for the j -th user in cell i and $z_{i,j} \sim \mathcal{CN}(0, N_0)$ is the corresponding additive white Gaussian complex noise. Under this model, the achieved SINR in downlink for the UT $_{i,j}$ is given by

$$\Lambda_{i,j}^{\text{DL}} = \frac{|\mathbf{w}_{i,j}^H \mathbf{h}_{i,i,j}|^2}{\sum_{(n,k) \neq (i,j)} |\mathbf{w}_{n,k}^H \mathbf{h}_{n,i,j}|^2 + N_0}. \quad (1)$$

The denominator terms of (1) represent the intra-cell interference, inter-cell interference and the noise (in order as they

appear). The downlink sum power minimization problem can be formulated as the following optimization problem given by²

$$\begin{aligned} \min_{\mathbf{w}_{i,j} \forall i,j} \quad & \sum_{i,j} \mathbf{w}_{i,j}^H \mathbf{w}_{i,j} \\ \text{s.t.} \quad & \Lambda_{i,j}^{\text{DL}} \geq \gamma_{i,j} \quad \forall i,j. \end{aligned} \quad (2)$$

B. Algorithm Design

As show in [18], the optimization problem (2) can be reformulated as a second order conic programming (SOCP) problem and, strong duality holds for this problem. Following the approach of [9], we solve this problem using duality theory. Accordingly, we introduce the Lagrange multiplier $\frac{\lambda_{i,j}}{N_t}$ associated with the downlink SINR constraints. The Lagrangian is given by

$$\begin{aligned} L(\mathbf{w}, \boldsymbol{\lambda}) = & \sum_{i,j} \mathbf{w}_{i,j}^H \mathbf{w}_{i,j} - \sum_{i,j} \frac{\lambda_{i,j}}{N_t} \left[\frac{|\mathbf{w}_{i,j}^H \mathbf{h}_{i,i,j}|^2}{\gamma_{i,j}} \right. \\ & \left. - \sum_{(n,k) \neq (i,j)} |\mathbf{w}_{n,k} \mathbf{h}_{n,i,j}|^2 - N_0 \right]. \end{aligned} \quad (3)$$

Note that the Lagrange multiplier is scaled by the factor N_t in order to ensure that the sum power in the system is finite, when the dimensions of the system grow large (in terms of the number of antennas on the BS and number of UTs). In fact, the Lagrange multipliers $\frac{\lambda_{i,j}}{N_t}$ can be interpreted as the dual uplink powers in the formulation of the dual uplink problem obtained in the following manner.

Rearranging (3), we obtain

$$L(\mathbf{w}, \boldsymbol{\lambda}) = \sum_{i,j} \frac{\lambda_{i,j} N_0}{N_t} - \sum_{i,j} \mathbf{w}_{i,j}^H \mathbf{B}_{i,j} \mathbf{w}_{i,j}, \quad (4)$$

where the matrix $\mathbf{B}_{i,j}$ is given by

$$\begin{aligned} \mathbf{B}_{i,j} = & \mathbf{I} - \left(1 + \frac{1}{\gamma_{i,j}} \right) \frac{\lambda_{i,j}}{N_t} \mathbf{h}_{i,i,j} \mathbf{h}_{i,i,j}^H \\ & + \sum_{n,k} \frac{\lambda_{n,k}}{N_t} \mathbf{h}_{i,n,k} \mathbf{h}_{i,n,k}^H \\ = & \mathbf{I} - \frac{\lambda_{i,j}}{\gamma_{i,j} N_t} \mathbf{h}_{i,i,j} \mathbf{h}_{i,i,j}^H + \sum_{(n,k) \neq (i,j)} \frac{\lambda_{n,k}}{N_t} \mathbf{h}_{i,n,k} \mathbf{h}_{i,n,k}^H. \end{aligned} \quad (5)$$

The dual uplink problem corresponding to the optimization in (2) is formulated as

$$\begin{aligned} \min_{\lambda_{i,j}, \forall i,j} \quad & \sum_{i,j} \frac{\lambda_{i,j}}{N_t} N_0 \\ \text{s.t.} \quad & \Lambda_{i,j}^{\text{UL}} \geq \gamma_{i,j}, \quad \forall i,j \end{aligned} \quad (6)$$

where the left hand side of the constraint equation represents the uplink SINR given by

$$\Lambda_{i,j}^{\text{UL}} = \frac{\frac{\lambda_{i,j}}{N_t} |\hat{\mathbf{w}}_{i,j}^H \mathbf{h}_{i,i,j}|^2}{\sum_{(n,k) \neq (i,j)} \frac{\lambda_{n,k}}{N_t} |\hat{\mathbf{w}}_{i,j}^H \mathbf{h}_{i,n,k}|^2 + \|\hat{\mathbf{w}}_{i,j}\|_2^2}$$

where $\hat{\mathbf{w}}_{i,j}$ denotes the corresponding uplink receive filter.

²In order to keep the analysis simple, in this work we do not consider the issue of max power constraint per BS.

We now provide a brief description of the beamforming algorithm presented in [9]. Before introducing the algorithm, we define the following matrices.

$$\begin{aligned}\mathbf{H}_{i,n} &= [\mathbf{h}_{i,n,1}, \dots, \mathbf{h}_{i,n,K}] \in \mathbb{C}^{N_t \times K} \\ \mathbf{H}_i &= [\mathbf{H}_{i,1}, \dots, \mathbf{H}_{i,N}] \in \mathbb{C}^{N_t \times NK} \\ \boldsymbol{\lambda}_i &= [\lambda_{i,1}, \dots, \lambda_{i,K}] \in \mathbb{C}^{K \times 1}, \\ \boldsymbol{\Lambda} &= \text{diag}[\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_N] \in \mathbb{C}^{NK \times NK}.\end{aligned}$$

We also define the matrix $\boldsymbol{\Sigma}_i^\lambda = \frac{1}{N_t} \mathbf{H}_i \boldsymbol{\Lambda} \mathbf{H}_i^H \in \mathbb{C}^{N_t \times N_t}$.

Algorithm 1 (Centralized Algorithm - **CBF**). *Perform the following steps.*

- Starting from any initial $\lambda_{i,j}^0 > 0 \forall i, j$ the uplink power allocation is given by $\lambda_{i,j} \triangleq \lim_{t \rightarrow \infty} \lambda_{i,j}^t$, where

$$\lambda_{i,j}^{t+1} = \frac{1}{\frac{1}{N_t} \left(1 + \frac{1}{\gamma_{i,j}}\right) \mathbf{h}_{i,i,j}^H (\boldsymbol{\Sigma}_i^{\lambda^t} + \mathbf{I}_{N_t})^{-1} \mathbf{h}_{i,i,j}} \quad \forall i, j \quad (7)$$

where $\boldsymbol{\Sigma}_i^{\lambda^t} = \frac{1}{N_t} \mathbf{H}_i \boldsymbol{\Lambda}^t \mathbf{H}_i^H$ and $\boldsymbol{\Lambda}^t = \text{diag}[\boldsymbol{\lambda}_1^t, \dots, \boldsymbol{\lambda}_N^t]$.

- The optimal receive uplink receive filter is given by

$$\hat{\mathbf{w}}_{i,j} = \frac{1}{\sqrt{N_t}} \left(\sum_{n,k} \frac{\lambda_{n,k}}{N_t} \mathbf{h}_{i,n,k} \mathbf{h}_{i,n,k}^H + \mathbf{I} \right)^{-1} \mathbf{h}_{i,i,j}. \quad (8)$$

- The optimal transmit downlink beamforming vectors are given by $\mathbf{w}_{i,j} = \sqrt{\frac{\delta_{i,j}}{N_t}} \hat{\mathbf{w}}_{i,j}$, where $\delta_{i,j}$ is given as

$$\boldsymbol{\delta} = \mathbf{F}^{-1} \mathbf{1}_{N_0}.$$

Here,

$$\begin{aligned}\boldsymbol{\delta}_i &= [\delta_{i,1}, \dots, \delta_{i,K}] \in \mathbb{C}^{K \times 1} \\ \boldsymbol{\delta} &= [\boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_N] \in \mathbb{C}^{NK \times 1} \\ \mathbf{1} &\in [1, \dots, 1]^T \in \mathbb{R}^{NK \times 1}\end{aligned}$$

and the elements of the matrix $\mathbf{F} \in \mathbb{C}^{NK \times NK}$ and the submatrix $\mathbf{F}^{i,j} \in \mathbb{C}^{K \times K}$ are given by,

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}^{1,1} & \dots & \mathbf{F}^{1,N} \\ \vdots & \ddots & \vdots \\ \mathbf{F}^{N,1} & \dots & \mathbf{F}^{N,N} \end{pmatrix} \quad (9)$$

$$\mathbf{F}_{j,k}^{i,n} \triangleq \begin{cases} \frac{1}{\gamma_{i,j} N_t} |\hat{\mathbf{w}}_{i,j}^H \mathbf{h}_{i,i,j}|^2, & n = i, k = j \\ \frac{1}{N_t} |\hat{\mathbf{w}}_{n,k}^H \mathbf{h}_{n,i,j}|^2, & (n, k) \neq (i, j). \end{cases} \quad (10)$$

We remark that the scaling of $\sqrt{N_t}$ in the expression for the uplink receive filter (8), and in the definition of the scaling factor $\delta_{i,j}$ ensure that the total power in the system is finite, when the dimensions of the system grow large.

As mentioned before, the solution provided in [9] cannot be implemented in a distributed manner. The computation of dual uplink power ($\lambda_{i,j}$) and the scaling factors ($\delta_{i,j}$) requires a central station which has the global CSI knowledge. In what follows, we overcome this problem.

We will now formulate our *reduced overhead* beamforming algorithm. The main idea behind this algorithm is that under

the massive MIMO regime (i.e. when N_t and K become large), the parameters in (7) and (10) can be approximated by their asymptotic equivalents using results from RMT. Moreover, the computation of these parameters will only depend on the second order statistics (path-loss), and not on the fast fading component of the channel vectors. However, note that RMT results are not directly applicable to this scenario. This is due to the fact that the computation of $\lambda_{i,j}$ in (7) explicitly depend on the channel vectors (RMT results require that the matrix $\boldsymbol{\Lambda}$ in (7) are independent of the channel matrices). This imposes additional technical difficulties in the application of RMT results in our scenario. Our approach in this work is to first propose an algorithm that depends only on the second order statistics of the channel vectors (the path-loss in our case). Mathematically/theoretically, it is not ensured that it achieves the optimal solution. However, we apply such an algorithm to the original system set-up, and prove that this algorithm is optimal in the large system domain.

We will hereby represent the dual uplink power, the uplink and downlink beamforming vectors of the decentralized algorithm by the notation $\mu_{i,j}$, $\hat{\mathbf{v}}_{i,j}$ and $\mathbf{v}_{i,j}$, respectively, which are the counterparts of $\lambda_{i,j}$, $\hat{\mathbf{w}}_{i,j}$ and $\mathbf{w}_{i,j}$ of the CBF algorithm.

Algorithm 2 (Reduced Overhead Beamforming algorithm - **ROBF**). *Perform the following steps.*

- Starting from any initial $\mu_{i,j}^0 > 0 \forall i, j$ the uplink power allocation is given by $\mu_{i,j} \triangleq \lim_{t \rightarrow \infty} \mu_{i,j}^t$, where

$$\mu_{i,j}^{t+1} = \frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i^t} \quad \forall i, j \quad (11)$$

and \bar{m}_i^t is evaluated as $\bar{m}_i^t \triangleq \lim_{p \rightarrow \infty} \bar{m}_i^{t,p}$ (initializing with any $\bar{m}_i^{t,0} > 0, \forall i$)

$$\bar{m}_i^{t,p} = \left(\frac{1}{N_t} \sum_{n=1}^N \sum_{k=1}^K \frac{\sigma_{i,n,k} \mu_{n,k}^t}{1 + \sigma_{i,n,k} \mu_{n,k}^t \bar{m}_i^{t,p-1}} + 1 \right)^{-1}. \quad (12)$$

- The optimal receive uplink receive filter is given by

$$\hat{\mathbf{v}}_{i,j} = \sqrt{\frac{1}{N_t}} \left(\sum_{n,k} \frac{\mu_{n,k}}{N_t} \mathbf{h}_{i,n,k} \mathbf{h}_{i,n,k}^H + \mathbf{I} \right)^{-1} \mathbf{h}_{i,i,j}. \quad (13)$$

- The optimal transmit downlink beamforming vectors are given by $\mathbf{v}_{i,j} = \sqrt{\frac{\bar{\delta}_{i,j}}{N_t}} \hat{\mathbf{v}}_{i,j}$. The scaling factor $\bar{\delta}_{i,j}$ is given as

$$\bar{\boldsymbol{\delta}} = (\mathbf{I} - \boldsymbol{\Gamma} \boldsymbol{\Delta})^{-1} \boldsymbol{\rho}, \quad (14)$$

where

$$\begin{aligned}\bar{\boldsymbol{\delta}}_i &= [\bar{\delta}_{i,1}, \bar{\delta}_{i,2}, \dots, \bar{\delta}_{i,K}]^T \in \mathbb{R}^{K \times 1} \\ \bar{\boldsymbol{\delta}} &= [\bar{\boldsymbol{\delta}}_1, \bar{\boldsymbol{\delta}}_2, \dots, \bar{\boldsymbol{\delta}}_N]^T \in \mathbb{R}^{NK \times 1} \\ \boldsymbol{\gamma}_i &= \left[\frac{\gamma_{i,1}}{\sigma_{i,i,1} \bar{G}_{i,i,1} \bar{m}_i^2}, \dots, \frac{\gamma_{i,K}}{\sigma_{i,i,K} \bar{G}_{i,i,K} \bar{m}_i^2} \right]^T \\ \boldsymbol{\gamma} &= [\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_N]^T \\ \boldsymbol{\Gamma} &= \text{diag}(\boldsymbol{\gamma})\end{aligned}$$

and the matrix $\Delta \in \mathbb{C}^{NK \times NK}$ is defined as

$$\Delta = \begin{pmatrix} \Delta^{1,1} & \dots & \Delta^{1,N} \\ \vdots & \ddots & \vdots \\ \Delta^{N,1} & \dots & \Delta^{N,N} \end{pmatrix} \quad (15)$$

where each submatrix $\Delta^{i,j} \in \mathbb{C}^{K \times K}$ is given by

$$\Delta_{j,k}^{i,n} \triangleq \begin{cases} 0, & n = i, k = j \\ \frac{1}{N_t} \bar{G}_{n,i,j} \bar{G}_{n,n,k} \bar{m}'_n, & (n,k) \neq (i,j). \end{cases} \quad (16)$$

\bar{m}'_i can be evaluated from \bar{m}_i as

$$\bar{m}'_i = \frac{\bar{m}_i^2}{1 - \frac{1}{N_t} \sum_{n=1}^N \sum_{k=1}^K \frac{(\sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}} \quad (17)$$

and the term

$$\bar{G}_{i,n,k} = \frac{\sigma_{i,n,k}}{(1 + \mu_{n,k} \sigma_{i,n,k} \bar{m}_i)^2}. \quad (18)$$

The vector $\rho_i = \left[\frac{N_0}{\sigma_{i,i,1} \bar{G}_{i,i,1} \bar{m}_i^2}, \dots, \frac{N_0}{\sigma_{i,i,K} \bar{G}_{i,i,K} \bar{m}_i^2} \right]^T$ and $\rho = [\rho_1, \rho_2, \dots, \rho_N]^T \in \mathbb{R}^{NK \times 1}$.

We now provide some remarks on this algorithm.

Lemma 1. *The iterative algorithm (11) converges to a fixed point.*

Proof. The proof is provided in Appendix B. \square

In section III, we characterize the solution provided by this fixed point, and show that it is asymptotically optimal in the sense that when the dimensions of the system grow large, the achieved uplink and downlink SINR by this algorithm exactly match the target SINR, and the allocated uplink and downlink powers match the optimal solution.

In the rest of the paper, we address the decentralized beamforming algorithm by the acronym ROBF. We now discuss the practical advantages of the ROBF algorithm over the CBF algorithm.

C. ROBF Algorithm - Signaling Overhead and Complexity Reduction

We now provide a brief discussion on the signaling overhead and complexity reduction associated with the ROBF algorithm.

Signaling Overhead:

Recall that in the ROBF algorithm only the statistical information (path loss) must be exchanged between the BSs where as in the CBF algorithm, the fast fading co-efficients need to be exchanged. Let us denote the channel coherence time by τ_{coh} units, and the long term time constant by τ_{LT} units (over which the statistical properties of the channel change). It has been demonstrated in some prior works related to channel measurements, that in an urban setting the statistical information of the channel can be viewed as constant roughly for over 100 channel coherence intervals [19].

We now characterize the signaling overhead for the two algorithms. For the CBF algorithm, exchanging full CSI would imply NN_tK complex channel coefficients or $2NN_tK$ real numbers every τ_{coh} units of time. This would amount to

exchanging NN_tK complex channel coefficients or $2NN_tK$ real co-efficients. Therefore, the rate of information exchange for CBF algorithm would be

$$R_{\text{CBF}} = \frac{2NN_tK}{\tau_{\text{coh}}} \text{ real coefficients/sec.}$$

For the ROBF algorithm, NK real numbers (path loss information assuming i.i.d. channel model) must be exchanged every τ_{LT} units of time, and the resulting rate of information exchange is given by

$$R_{\text{ROBF}} = \frac{NN_t}{\tau_{\text{LT}}} \text{ real coefficients/sec.}$$

In order to get real feel of these numbers, assume $N = 3, N_t = 100, K = 40, \tau_{\text{LT}} = 22.6s, \tau_{\text{coh}} = 180ms$ [19]. Therefore,

$$R_{\text{CBF}} = 1.3 \times 10^5 \text{ real coefficients/sec;} \\ R_{\text{ROBF}} = 5.3 \text{ real coefficients/sec,}$$

Therefore, the ratio of two quantities can be given by

$$\frac{R_{\text{CBF}}}{R_{\text{ROBF}}} = 2N_t \left(\frac{\tau_{\text{LT}}}{\tau_{\text{coh}}} \right) \approx 10^4.$$

However, for correlated channel model, $R_{\text{ROBF}} = \frac{NN_tK}{\tau_{\text{LT}}}$ real coefficients/sec. Therefore

$$\frac{R_{\text{CBF}}}{R_{\text{ROBF}}} = 2 \left(\frac{\tau_{\text{LT}}}{\tau_{\text{coh}}} \right) \approx 100.$$

Finally, note that the exact number of bits to be exchanged depends on how these channel co-efficients are quantized (which is beyond the scope of this paper).

Implementation Complexity:

An exact characterization of the complexity associated with the ROBF algorithm is beyond the scope of this work. Nevertheless, we provide a brief analysis of the same.

Computing the uplink power allocation

Recall that implementing the iterations for the computation of the uplink power in the CBF algorithm requires matrix inversion operations, where as the ROBF algorithm only requires performing scalar operations. This tremendously reduces the computational complexity with respect to the ROBF algorithm (the complexity of inverting a matrix is provided next). Moreover, in the CBF algorithm the uplink power must be evaluated for every channel realization (fast fading CSI), where as the ROBF algorithm requires parameters to be computed only once (at the time scale of changing of slow fading CSI). Therefore, there is a huge reduction in the computational complexity.

Formulation of the downlink beamforming vectors

Note that the downlink beamforming vector in the ROBF algorithm (and also the CBF algorithm) is in the form of a regularized zero forcing (RZF) beamforming, which requires computation of a matrix inverse of dimension $N_t \times NK$. Therefore, its computational complexity scales as $N_t(NK)^2$. This can be computationally demanding especially in a massive MIMO setting. Fortunately, alternate schemes are being developed to implement RZF based on truncated polynomial expansion incur much less computational complexity

as compared to matrix inversion, which are shown to have performance very close to RZF beamforming [20].

The price to pay for the reduction in information exchange between the BSs is that in the ROBF algorithm, the target SINR values are not met perfectly for every channel realization. In fact, the achieved SINR in the downlink fluctuates around the target SINR. However, we will show through simulations that even for practical values of the system dimensions, the fluctuations of the achieved SINR around the target SINR value are small. In order to make sure that the target SINR requirements are satisfied, one could solve the optimization problem with ROBF algorithm by considering a higher value target SINR (than the actual desired one) in order to compensate for the fluctuations.

Next, we will show that the performance of ROBF algorithm perfectly matches the CBF algorithm when the number of antennas per BS and the number of UTs become large, i.e., in the regime of massive MIMO systems. We also provide simulation results to examine the performance of the ROBF algorithm.

III. ALGORITHM ANALYSIS

In this section, we provide extensive analysis of the ROBF algorithm. In the rest of the paper, we will use the phrase "large system" to refer to the regime when the number of antennas per BS and the number of UTs per-cell become large, i.e., $N_t, K \rightarrow \infty$ while their ratio $\frac{N_t}{K}$ tends to a finite constant $0 < \beta < \infty$, as considered in some past works in this field [5]. However, we mention that our results provide tight approximations for practical system dimensions of massive MIMO systems. Specifically, we focus on the following aspects:

- We first characterize a lower bound on the feasible SINR targets for the ROBF algorithm.
- We prove that the uplink and the downlink SINR achieved by the ROBF algorithm asymptotically converge to their target value in the large system regime.
- We prove that the uplink and the downlink power allocations yielded by the ROBF algorithm asymptotically converge to the respective values of the CBF algorithm, hence making the ROBF algorithm optimal in the large system regime.

We first start with the characterization of feasible SINR targets for the ROBF algorithm.

A. Feasible SINR targets for the ROBF algorithm

Throughout this subsection, we will use the notations γ to denote the vector of target SINRs as follows:

$$\gamma_i = [\gamma_{i,1}, \dots, \gamma_{i,K}]; \quad \gamma = [\gamma_1, \dots, \gamma_N].$$

Similarly, we define the vectors

$$\lambda_i = [\lambda_{i,1}, \dots, \lambda_{i,K}]; \quad \lambda = [\lambda_1, \dots, \lambda_N],$$

$$\mu_i = [\mu_{i,1}, \dots, \mu_{i,K}]; \quad \mu = [\mu_1, \dots, \mu_N].$$

We will now define the notion of feasible SINR target for the ROBF algorithm. Feasible SINR target implies that the following three conditions must be satisfied :

[C1] The iterations of the fixed point equation (11) must converge to a finite value μ . This implies that given a target SINR vector γ , there exists $\mu < \infty$ that satisfies the fixed point equation

$$\gamma_{i,j} = \sigma_{i,i,j} \mu_{i,j} \bar{m}_i \quad \forall i, j. \quad (19)$$

Further from the property of the fixed point equation, for a given γ there exists a unique μ satisfying (19) (if the target SINR vector is feasible). Let us define such a pair of vector by $\{\gamma, \mu\}$.

[C2] For every pair of vectors $\{\gamma, \mu\}$ that satisfy (19), the matrix $\mathbf{I} - \Gamma\Delta$ must be invertible.

[C3] The elements of the vector $(\mathbf{I} - \Gamma\Delta)^{-1} \rho$ must be positive (in order for $\sqrt{\frac{\delta_{i,j}}{N_t}}$ to be real).

We focus on the conditions [C1] and [C2] and defer [C3] to the end of the section. We first state the main result of this subsection and later on provide the proof.

Theorem 1. *Every target SINR vector γ whose elements satisfy the condition*

$$\frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\frac{\sigma_{max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}{1 + \frac{\sigma_{max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}} < 1. \quad (20)$$

are feasible for the ROBF algorithm, where $\sigma_{max}(n) =$

We remark that in our result, the condition (20) is only a sufficient condition. The proof of this theorem involves several steps which will be illustrated in the rest of this subsection. We proceed as follows.

First, we establish a relationship between the two conditions [C1] and [C2].

Lemma 2. *For every pair of vectors $\{\gamma, \mu\}$ that satisfy (19), the matrix $\mathbf{I} - \Gamma\Delta$ is invertible.*

Proof. The proof can be found in Appendix C, part I. \square

Lemma 2 implies that the condition [C2] is automatically satisfied for every pair of vectors $\{\gamma, \mu\}$ that satisfy the condition [C1]. Following the result of Lemma 2, we consider characterizing only the set of target SINR vectors γ that satisfy the condition [C1].

The exact set of feasible target SINR vectors satisfying the condition [C1] is difficult to be characterized in closed form. Therefore, we only establish a lower bound on this set. In order to do so, we consider a modified system in which the inter-cell interference path loss coefficients are replaced by

$$\sigma_{max}(n) \triangleq \sup_k \sigma_{i,n,k}, \quad \forall n \neq i. \quad (21)$$

where $i = 1, \dots, N$. Therefore, in the modified system $\sigma_{i,n,k} = \sigma_{max}(n)$.

Let us consider the ROBF algorithm applied to both the original and the modified systems. The fixed point equation for the computation of the uplink power allocation for the original system must satisfy the following equations (we use

the superscript "org" to represent the original system):

$$\begin{aligned}\gamma_{i,j} &= \sigma_{i,i,j} \mu_{i,j}^{\text{org}} \bar{m}_i^{\text{org}} \quad \forall i, j, \\ \frac{1}{\bar{m}_i^{\text{org}}} &= \frac{1}{N_t} \sum_{k=1}^K \frac{\sigma_{i,i,k} \mu_{i,k}^{\text{org}}}{1 + \sigma_{i,i,k} \mu_{i,k}^{\text{org}} \bar{m}_i^{\text{org}}} \\ &+ \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\sigma_{i,n,k} \mu_{n,k}^{\text{org}}}{1 + \sigma_{i,n,k} \mu_{n,k}^{\text{org}} \bar{m}_i^{\text{org}}} + 1.\end{aligned}\quad (22)$$

Similarly, the fixed point equation for the *modified system* must satisfy the following equations (we use the superscript "mod" to represent the modified system):

$$\begin{aligned}\gamma_{i,j} &= \sigma_{i,i,j} \mu_{i,k}^{\text{mod}} \bar{m}_i^{\text{mod}} \\ \frac{1}{\bar{m}_i^{\text{mod}}} &= \frac{1}{N_t} \frac{\sigma_{i,i,k} \mu_{i,k}^{\text{mod}}}{1 + \sigma_{i,i,k} \mu_{i,k}^{\text{mod}} \bar{m}_i^{\text{mod}}} \\ &+ \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\sigma_{max}(n) \mu_{n,k}^{\text{mod}}}{1 + \sigma_{max}(n) \mu_{n,k}^{\text{mod}} \bar{m}_i^{\text{mod}}} + 1.\end{aligned}\quad (24)$$

In what follows, we will consider the ROBF algorithm applied to the *modified system* and characterize the feasible SINR targets corresponding to this system. The set of feasible target SINR of the ROBF algorithm applied to the *modified system* will act as a lower bound on the set of feasible target SINR of the ROBF algorithm applied to the original system. Intuitively, this is not hard to see. In the *modified system*, the path losses corresponding to the inter-cell interference links are scaled up to $\sigma_{max}(n)$. Therefore, the *modified system* represents a more interference limited regime as compared to the original system. Hence, any SINR feasible for ROBF applied to the *modified system* should be feasible for the ROBF algorithm applied to the original system as well. We will later on make rigorous arguments to prove that the above statement in Lemma 3.

Proposition 2. *The set of feasible SINR targets for the ROBF algorithm applied to the modified system must satisfy for $i = 1, \dots, N$,*

$$\frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\frac{\sigma_{max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}{1 + \frac{\sigma_{max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}} < 1. \quad (26)$$

Proof. The proof is provided in Appendix C, part II. \square

We will now present the following Corollary.

Corollary 1:

In the large system regime, the set of feasible SINR targets for the ROBF algorithm applied to the modified system must satisfy for $i = 1, \dots, N$,

$$\limsup_{N_t, K \rightarrow \infty} \left[\frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\frac{\sigma_{max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}{1 + \frac{\sigma_{max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}} \right] < 1. \quad (27)$$

The arguments for the Corollary is also provided in Appendix C, part II.

Finally, we will show that the feasibility condition of (26) will act as a lower bound on the set of feasible SINR targets for the ROBF algorithm applied to the original system.

Lemma 3. *Any target SINR feasible for the ROBF algorithm applied to the modified system is feasible for the ROBF algorithm applied to the original system. Hence the feasible SINR targets for the modified system will act as a lower bound for the feasible SINR target of the original system.*

Proof. The proof is provided in Appendix C, part III. \square

Lastly, we establish the condition [C3]. This is rather straight forward and can be seen by the following steps: From Appendix C, we have established that for target SINR values that satisfy the condition in Proposition 2, $\rho(\Gamma \Delta) < 1$. Therefore, using series expansion for the matrix $(\mathbf{I} - \Gamma \Delta)^{-1}$ [21], we have

$$(\mathbf{I} - \Gamma \Delta)^{-1} \rho = \left(\sum_{j=1}^{\infty} (\Gamma \Delta)^j \right) \rho. \quad (28)$$

Since the elements of $\Gamma \Delta$ and ρ are positive, their sum will also be positive. Therefore, the condition [C3] holds true.

We end this section by providing the feasibility conditions in two special cases in which the feasibility conditions are both necessary and sufficient:

Example 1: Single Cell Case

For the isolated single cell case, (27) reduces to the following

$$\frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_k}{1 + \gamma_k} < 1. \quad (29)$$

When all the UTs are demanding the same SINR $\gamma_k = \gamma \forall k$, the set of feasible SINR is given by $\gamma < \left(\frac{K}{\min\{N_t, K\}} - 1 \right)^{-1}$. The result can be interpreted as follows. In the case of a single cell, as long as $N_t \geq K$, any finite SINR target is supportable by the ROBF algorithm (for the case when all $\gamma_{i,j}$ are equal). In other words, when $N_t \geq K$, the BS has enough degrees of freedom to sever all the UTs in the system. Note that (29) matches the feasibility conditions derived in [18] for the case of a single cell system.

Example 2: 2-Cell Wyner Model

Consider a perfectly symmetric multi-cell system in which the path loss for the intra-cell links are equal to 1, and the path loss of the inter-cell links are equal to ϵ . Every UT demands the same SINR target given by γ . In this case, (27) reduces to

$$\frac{K}{N_t} \left(\frac{\gamma}{1 + \gamma} + \frac{\epsilon \gamma}{1 + \epsilon \gamma} \right) < 1. \quad (30)$$

This condition allows us to examine the dependency of the feasible SINR target on various system parameters. In particular, we examine the dependency of the feasible γ on N_t by a simple numerical example. We plot the downlink transmit power (obtained by running the ROBF algorithm) as a function of γ for different values of N_t in Figure 1. We consider $K = 50$ UTs per cell and $\epsilon = 0.5$. It can be seen that beyond a certain cut off value of the target SINR, the downlink power grows unbounded. This is precisely the value of the target

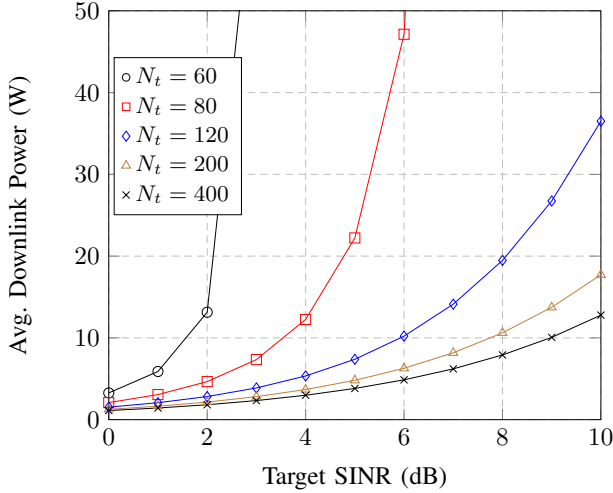


Fig. 1. Downlink power Vs target SINR for $K = 50$ UTs per cell.

SINR at which the ROBF algorithm becomes infeasible. It can be verified that the cut off value of γ is the one for which the condition in (30) is not satisfied. Finally, note that higher the number of transmit antennas per BS, higher is the cut off value of the target SINR. Once again, this is due to the availability of greater number of spatial degrees of freedom. In this cases when $N_t \geq 2K$, any finite target SINR is achievable (note that $2K$ is the total number of UTs in two cells). The feasibility conditions for the special case of two cell Wyner model was derived in [11] as well.

B. Convergence of the Uplink and Downlink SINR

Next we focus on the achieved SINR in the uplink and downlink for the ROBF algorithm. We first start with the analysis of the uplink of the ROBF algorithm. Before we proceed, we make the following observations. Recall that the achieved SINR in the uplink for the ROBF algorithm is given by

$$\Lambda_{i,j}^{\text{UL}}(\boldsymbol{\mu}) = \frac{\frac{\mu_{i,j}}{N_t} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,i,j}|^2}{\sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k}}{N_t} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,n,k}|^2 + \|\hat{\mathbf{v}}_{i,j}\|_2^2}}. \quad (31)$$

Note that in our notation $\Lambda_{i,j}^{\text{UL}}(\boldsymbol{\mu})$, we have explicitly mentioned the achieved uplink SINR as a function of the parameters of the ROBF algorithm $\boldsymbol{\mu}$ (slightly deviating from the notation for the uplink SINR introduced in Section II). Since the uplink receive filter in the ROBF algorithm (13) are minimum mean square error (MMSE) form, the expression for the achieved uplink SINR can be given in alternate form as

$$\Lambda_{i,j}^{\text{UL}}(\boldsymbol{\mu}) = \frac{\mu_{i,j}}{N_t} \mathbf{h}_{i,i,j}^H (\boldsymbol{\Sigma}_i^\mu + \mathbf{I}_{N_t})^{-1} \mathbf{h}_{i,i,j} \quad (32)$$

where $\boldsymbol{\Sigma}_i^\mu = \sum_{i,j} \frac{\mu_{i,j}}{N_t} \mathbf{h}_{i,i,j} \mathbf{h}_{i,i,j}^H$ and $\boldsymbol{\Sigma}_i^\mu = \boldsymbol{\Sigma}_i^\mu - \frac{\mu_{i,j}}{N_t} \mathbf{h}_{i,i,j} \mathbf{h}_{i,i,j}^H$. Similarly since the uplink receive filter (8) in the CBF algorithm are MMSE, the achieved uplink SINR for the CBF algorithm is given by

$$\Lambda_{i,j}^{\text{UL}}(\boldsymbol{\lambda}) = \frac{\lambda_{i,j}}{N_t} \mathbf{h}_{i,i,j}^H (\boldsymbol{\Sigma}_i^\lambda + \mathbf{I}_{N_t})^{-1} \mathbf{h}_{i,i,j} \quad \forall i, j. \quad (33)$$

Also, since the CBF algorithm is optimal,

$$\Lambda_{i,j}^{\text{UL}}(\boldsymbol{\lambda}) = \gamma_{i,j} \quad \forall i, j. \quad (34)$$

Also, recall that the downlink SINR for the ROBF algorithm is given by

$$\Lambda_{i,j}^{\text{DL}}(\boldsymbol{\mu}) = \frac{|\mathbf{v}_{i,j}^H \mathbf{h}_{i,i,j}|^2}{\sum_{k \neq j} |\mathbf{v}_{i,k}^H \mathbf{h}_{i,i,j}|^2 + \sum_{n \neq i,k} |\mathbf{v}_{n,k}^H \mathbf{h}_{n,i,j}|^2 + N_0}. \quad (35)$$

We now provide the following convergence result:

Theorem 3. *In the large system regime, the achieved uplink and downlink SINR for the ROBF algorithm converge almost surely to the target SINR $\gamma_{i,j}$. Mathematically stating*

$$\Lambda_{i,j}^{\text{UL}}(\boldsymbol{\mu}) \xrightarrow[N_t, K \rightarrow \infty]{\text{a.s.}} \gamma_{i,j} \quad \forall i, j, \quad (36)$$

$$\Lambda_{i,j}^{\text{DL}}(\boldsymbol{\mu}) \xrightarrow[N_t, K \rightarrow \infty]{\text{a.s.}} \gamma_{i,j} \quad \forall i, j. \quad (37)$$

Proof. The details of the proof for the convergence of the achieved uplink SINR (36) can be found in Appendix D.

Next we proceed to the convergence proof for the downlink SINR (37). The proof utilizes the following lemma on the convergence result for the downlink interference terms.

Lemma 4. *The downlink interference term corresponding to the intra-cell and inter-cell interference converge in the large system regime to the following:*

$$\sum_{(n,k) \neq (i,j)} |\mathbf{v}_{n,k}^H \mathbf{h}_{n,i,j}|^2 \asymp \sum_{(n,k) \neq (i,j)} \frac{\bar{\delta}_{n,k} \bar{G}_{n,i,j} \bar{G}_{n,n,k} \bar{m}'_n}{N_t}. \quad (38)$$

The lemma is proved in Appendix E. We now proceed to the convergence of the downlink SINR. Using Lemma 4 and Lemma 11, it can be concluded that the downlink SINR asymptotically converges to

$$\Lambda_{i,j}^{\text{DL}}(\boldsymbol{\mu}) \xrightarrow[N_t, K \rightarrow \infty]{\text{a.s.}} \frac{\bar{\delta}_{i,j} \sigma_{i,i,j} \bar{G}_{i,i,j} \bar{m}_i^2}{\frac{1}{N_t} \sum_{(n,k) \neq (i,j)} \bar{\delta}_{n,k} \bar{G}_{n,i,j} \bar{G}_{n,n,k} \bar{m}'_n + N_0}. \quad (39)$$

It can be easily verified from (14) that the right hand side of (39) is equal to the target SINR $\gamma_{i,j}$, thus completing the proof. \square

C. Asymptotic optimality of the Uplink and Downlink Power Allocation

We will now focus on the optimality of the uplink and downlink power allocation of the ROBF algorithm in the large system regime.

Consider the Lagrangian of the downlink minimization problem in its two forms as in equations (3) and (4). Our proof proceeds by plugging in the solution obtained by the ROBF algorithm $(\mathbf{v}, \boldsymbol{\mu})$ into the Lagrangian and examining its properties in the large system regime.

Lemma 5. *The following results hold true for the Lagrangian in the large system limit:*

$$\lim_{N_t, K \rightarrow \infty} L(\mathbf{v}, \boldsymbol{\mu}) = \lim_{N_t, K \rightarrow \infty} \sum_{i,j} \mathbf{v}_{i,j}^H \mathbf{v}_{i,j}, \quad (40)$$

$$\lim_{N_t, K \rightarrow \infty} L(\mathbf{v}, \boldsymbol{\mu}) = \lim_{N_t, K \rightarrow \infty} \sum_{i,j} \frac{\mu_{i,j} N_0}{N_t}. \quad (41)$$

Proof. Please refer to Appendix F. \square

We can draw the following inference from the result of Lemma 5.

Corollary 2: The uplink and downlink power allocations yielded by the ROBF algorithm are equal in the asymptotic limit:

$$\lim_{N_t, K \rightarrow \infty} \sum_{i,j} \mathbf{v}_{i,j}^H \mathbf{v}_{i,j} = \lim_{N_t, K \rightarrow \infty} \sum_{i,j} \frac{\mu_{i,j} N_0}{N_t}. \quad (42)$$

We know that the uplink power allocation of the CBF algorithm satisfies

$$\sum_{i,j} \frac{\lambda_{i,j} N_0}{N_t} = \min_{\mathbf{w}} \max_{\boldsymbol{\lambda}} L(\mathbf{w}, \boldsymbol{\lambda}) \quad (43)$$

If we show that in the large system regime

$$\lim_{N_t, K \rightarrow \infty} \sum_{i,j} \frac{\mu_{i,j} N_0}{N_t} = \lim_{N_t, K \rightarrow \infty} \min_{\mathbf{w}} \max_{\boldsymbol{\lambda}} L(\mathbf{w}, \boldsymbol{\lambda}), \quad (44)$$

and the duality gap is zero, then this implies that the solution provided by the ROBF is optimal in the asymptotic domain. In other words, the optimal downlink power (which is the solution of the primal problem) is equal to $\lim_{N_t, K \rightarrow \infty} \min_{\mathbf{w}} \max_{\boldsymbol{\lambda}} L(\mathbf{w}, \boldsymbol{\lambda})$. In order to do so, we will prove the following result:

Lemma 6. *In the large system regime, the sum of uplink power allocation of the ROBF algorithm converges to the sum of uplink power allocation of the CBF algorithm.*

$$\sum_{i,j} \frac{\mu_{i,j}}{N_t} \simeq \sum_{i,j} \frac{\lambda_{i,j}}{N_t} \quad \forall i, j. \quad (45)$$

Proof. The result is proved in Appendix G. \square

Consequently, the solution provided by the ROBF algorithm is optimal to $\lim_{N_t, K \rightarrow \infty} \max_{\mathbf{w}} \min_{\boldsymbol{\lambda}} L(\mathbf{w}, \boldsymbol{\lambda})$. Also, from the result of Corollary 2, it follows that the downlink power allocation of the ROBF algorithm is also optimal in the large system regime. This concludes the proof.

IV. IMPACT OF IMPERFECT CSI AND PILOT CONTAMINATION ON THE PERFORMANCE OF ROBF ALGORITHM

In this section, we investigate the impact of CSI estimation errors and pilot contamination on the performance of the ROBF algorithm. Throughout this section, we assume that the slow fading co-efficient (path loss information) can be accurately estimated at the BS (since they remain constant for a long period of time, they are easy to estimate, see

for e.g. [22]). Further, for the fast fading co-efficients, we assume reciprocity between uplink and downlink channels, and consider the time division duplexing (TDD) model of channel estimation, i.e., estimation via uplink pilots.

A comprehensive design of the optimal beamforming vectors in the presence of CSI estimation errors and pilot contamination issue is out of the scope of this paper. Alternately, we take the following approach: First, we assume that the BS treats the CSI estimate as the true CSI and implements the ROBF algorithm directly. In this case, we derive the asymptotic equivalent of the downlink SINR achieved by the ROBF algorithm. Using numerical results, we investigate the impact of CSI estimation errors and pilot contamination on the performance of the ROBF algorithm both in terms of the achieved SINR and downlink power. Then, exploiting the fact that the BS has the accurate knowledge of the slow fading co-efficient, we propose a heuristic adaptation of the ROBF algorithm in the presence of imperfect CSI and pilot contamination named as the modified ROBF (MROBF) algorithm. We show that under the massive MIMO regime, an algorithm in which parameters can be computed based on the channel statistics (rather than the fast fading CSI) such as the MROBF algorithm is more robust to CSI estimation and pilot contamination effects.

A. CSI Estimation

We now describe the CSI training phase for the estimation of the fast fading co-efficients. Let T be the length of the channel coherence interval, a part of which is dedicated for CSI estimation, and τ be the number of symbols used for pilots. Therefore, the UTs in every cell i transmit τ mutually orthogonal pilot symbols to their respective BSs during the training phase. We represent the pilot sequences used by the K UTs in each cell by the matrix $\sqrt{P_{\text{Tr}}} \boldsymbol{\Phi} \in \mathbb{C}^{\tau \times K}$, ($\tau \geq K$). The pilot sequences are repeated in each cell, hence leading to the issue of pilot contamination. The matrix $\boldsymbol{\Phi}$ satisfies $\boldsymbol{\Phi}^H \boldsymbol{\Phi} = \mathbf{I}_K$.

The signal received during the channel training phase denoted by $\mathbf{Y}_{i,\text{Tr}} \in \mathbb{C}^{N_t \times K}$ can be written as

$$\mathbf{Y}_{i,\text{Tr}} = \mathbf{H}_{i,i} \boldsymbol{\Phi}^T + \left(\sum_{n \neq i} \mathbf{H}_{i,n} \right) \boldsymbol{\Phi}^T + \mathbf{N}, \quad (46)$$

where $\mathbf{H}_{i,n} = [\mathbf{h}_{i,n,1}, \dots, \mathbf{h}_{i,n,K}]$ and $\mathbf{N} \in \mathbb{C}^{N_t \times \tau}$ with i.i.d. $\mathcal{CN}(0, 1)$ elements represents the noise during channel training phase. The MMSE estimate of $\mathbf{h}_{i,n,k}$ given $\mathbf{Y}_{i,\text{Tr}}$ can be given by [23]

$$\hat{\mathbf{h}}_{i,n,k} = \sigma'_{i,n,k} \left(\sum_{b=1}^N \mathbf{h}_{i,b,k} + \frac{\mathbf{n}_{\text{Tr}}}{\sqrt{P_{\text{Tr}}}} \right), \quad (47)$$

where $\sigma'_{i,n,k} = \sigma_{i,n,k} \left(\sum_{b=1}^N \sigma_{i,b,k} + \frac{1}{P_{\text{Tr}}} \right)^{-1}$. It can be verified that $\hat{\mathbf{h}}_{i,n,k}$ is distributed as

$$\hat{\mathbf{h}}_{i,n,k} \sim \mathcal{CN} \left(0, \sigma_{i,n,k}^2 \left(\sum_{b=1}^N \sigma_{i,b,k} + \frac{1}{P_{\text{Tr}}} \right)^{-1} \right). \quad (48)$$

For notational simplicity, let us denote $\hat{\sigma}_{i,n,k} = \sigma_{i,n,k}^2 \left(\sum_{b=1}^N \sigma_{i,b,k} + \frac{1}{P_{\text{tr}}} \right)^{-1}$. Note that

$$\hat{\sigma}_{i,n,k} = \sigma'_{i,n,k} \sigma_{i,n,k}. \quad (49)$$

ROBF Algorithm With CSI Estimates

Throughout this subsection, we assume that the BSs assume the CSI estimates to the true channel vector, and implement the ROBF algorithm. Note however that the computation of $\mu_{i,j}$ and $\bar{\delta}_{i,j}$ in (11) and (14) do not require the estimates of the fast fading co-efficient, and hence they can be implemented directly. The uplink receive filter with imperfect CSI denoted by $\hat{\mathbf{v}}_{i,j}^{\text{est}}$ can be formulated as

$$\hat{\mathbf{v}}_{i,j}^{\text{est}} = \sqrt{\frac{1}{N_t}} \Psi_i^{-1} \hat{\mathbf{h}}_{i,i,j} \quad (50)$$

where

$$\Psi_i = \sum_{n,k} \frac{\mu_{i,k}}{N_t} \hat{\mathbf{h}}_{i,n,k} \hat{\mathbf{h}}_{i,n,k}^H + \mathbf{I}_{N_t}. \quad (51)$$

The downlink beamforming vector denoted by $\mathbf{v}_{i,j}^{\text{est}}$ can be then computed as

$$\mathbf{v}_{i,j}^{\text{est}} = \sqrt{\frac{\bar{\delta}_{i,j}}{N_t}} \hat{\mathbf{v}}_{i,j}^{\text{est}}. \quad (52)$$

We now investigate the achieved SINR in the downlink under the ROBF algorithm with imperfect CSI. The expression for the downlink SINR can be given as in (35), by replacing $\mathbf{v}_{i,j}$ with $\hat{\mathbf{v}}_{i,j}^{\text{est}}$. The asymptotic equivalent of the downlink SINR can be obtained by analyzing it in the large system regime.

Theorem 4. *The achieved SINR in the downlink by the ROBF algorithm in the presence of imperfect CSI converges almost surely to the right hand side of (56) in the large system regime, where the term $\hat{G}_{n,n,k}$ is defined as*

$$\hat{G}_{n,n,k} = \frac{1}{1 + \xi_{n,k} \hat{\sigma}_{n,n,k} \bar{m}_n^{\text{est}}} \text{ and } \xi_{i,k} = \frac{\sum_{n=1}^N \sigma_{i,n,k} \mu_{n,k}}{\sigma_{i,i,k}}. \quad (53)$$

Further, \bar{m}_n^{est} can be computed as the solution to the fixed point equation

$$\bar{m}_n^{\text{est}} = \left(\frac{1}{N_t} \sum_{k=1}^K \frac{\xi_{n,k} \hat{\sigma}_{n,n,k}}{1 + \xi_{n,k} \hat{\sigma}_{n,n,k} \bar{m}_n^{\text{est}}} + 1 \right)^{-1}, \quad (54)$$

and $(\bar{m}'_n)^{\text{est}}$ can be computed from \bar{m}_n^{est} as

$$(\bar{m}'_n)^{\text{est}} = \frac{(\bar{m}_n^{\text{est}})^2}{1 - \frac{1}{N_t} \sum_{k=1}^K \frac{(\hat{\sigma}_{n,n,k} \xi_{n,k} \bar{m}_n^{\text{est}})^2}{(1 + \hat{\sigma}_{n,n,k} \xi_{n,k} \bar{m}_n^{\text{est}})^2}}. \quad (55)$$

Proof. The proof is provided in Appendix H. \square

B. Modified ROBF (MROBF) Algorithm

In this subsection, we propose a heuristic adaptation of the ROBF algorithm addressed as Modified ROBF (MROBF) algorithm which is designed to accommodate the effects of imperfect CSI and pilot contamination.

First, note that naive application of the ROBF algorithm may not yield good performance in the presence of imperfect CSI and pilot contamination. The following reasoning provides an intuitive understanding for this: Consider the interference arising from the signal of $UT_{n,k}$, ($n \neq i, k \neq j$) at the $UT_{i,j}$ (which does not use the same pilot as $UT_{i,j}$), i.e., the term $|\mathbf{w}_{n,k} \mathbf{h}_{n,i,j}|^2$. Following the derivation of Appendix H, it can be verified that this interference terms converges to

$$|\mathbf{w}_{n,k} \mathbf{h}_{n,i,j}|^2 \asymp \frac{\bar{\delta}_{n,k}}{N_t} \hat{\sigma}_{n,n,k} \hat{G}_{n,n,k}^2 (\bar{m}'_n)^{\text{est}} B_{n,k}. \quad (60)$$

Now consider the interference arising from the signal of $UT_{n,j}$, ($n \neq i$) at the $UT_{i,j}$ (which reuses the same pilot as $UT_{i,j}$), i.e., the term $|\mathbf{w}_{n,j} \mathbf{h}_{n,i,j}|^2$. It can be verified that asymptotically, this interference terms converges to

$$|\mathbf{w}_{n,j} \mathbf{h}_{n,i,j}|^2 \asymp \bar{\delta}_{n,j} \left(\sigma_{n,i,j} \sigma'_{n,n,j} \hat{G}_{n,n,j} \bar{m}_n^{\text{est}} \right)^2. \quad (61)$$

We note that the two interference terms are significantly different. In particular, (60) is scaled by a factor of $\frac{1}{N_t}$, (and for large N_t , this has a very low value). However, the term in (61) is not scaled. This is due to the fact that the BS cannot distinguish between the channels of $UT_{i,j}$ and $UT_{n,j}$ due to the pilot contamination effect. This implies that naive application of the ROBF algorithm may significantly underestimate the interference arising out of the UTs that reuse the same pilots. The algorithm performance can be enhanced by carefully accounting for these issues. This is indeed the main intuition behind the MROBF algorithm.

In the MROBF algorithm, we consider that the BS does not alter the structure of the beamforming vector, i.e., the BS retains the RZF beamforming vector as in (52). Nevertheless, the performance gain can still be obtained by redesigning the uplink power allocation, and computation of $\bar{\delta}$. We redesign the uplink power allocation on similar lines as that of the ROBF algorithm.

First note that in the case of perfect CSI, the achieved uplink SINR in the asymptotic limit is given by $\sigma_{i,i,j} \mu_{i,j} \bar{m}_i$. Further, the uplink power allocation is chosen to satisfy $\gamma_{i,j} = \sigma_{i,i,j} \mu_{i,j} \bar{m}_i$, or

$$\mu_{i,j} = \frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i} \quad \forall i, j. \quad (62)$$

We use a similar argument for the computation of the uplink power allocation in the case of imperfect CSI. In order to do so, consider the uplink SINR with uplink receive filter formulated as in (50)

$$(\Lambda_{i,j}^{\text{UL}}(\boldsymbol{\mu}))_{\text{est}} = \frac{\frac{\mu_{i,j}}{N_t} |\hat{\mathbf{v}}_{i,j}^{\text{est}} \mathbf{h}_{i,i,j}|^2}{\sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k}}{N_t} |\hat{\mathbf{v}}_{i,j}^{\text{est}} \mathbf{h}_{i,n,k}|^2 + \|\hat{\mathbf{v}}_{i,j}^{\text{est}}\|_2^2}. \quad (63)$$

Let us examine the uplink SINR in the large system domain. Using the derivation similar to Appendix H, we can replace

$$(\Lambda_{i,j}^{\text{DL}}(\boldsymbol{\mu}))_{\text{est}} \asymp \frac{\bar{\delta}_{i,j}(\hat{\sigma}_{i,i,j}\hat{G}_{i,i,j}\bar{m}_i^{\text{est}})^2}{\sum_{n \neq i}^N \bar{\delta}_{n,j} \left(\sigma_{n,i,j} \sigma'_{n,n,j} \hat{G}_{n,n,j} \bar{m}_n^{\text{est}} \right)^2 + \frac{1}{N_t} \sum_{n=1}^N \sum_{k \neq j}^K \bar{\delta}_{n,k} \hat{\sigma}_{n,n,k} \hat{G}_{n,n,k}^2 (\bar{m}'_n)^{\text{est}} B_{n,k}}, \quad (56)$$

where

$$B_{n,k} = \sigma_{n,i,j} + \hat{\sigma}_{n,n,j} (\xi_{n,j} \sigma_{n,i,j} \sigma'_{n,n,j} \hat{G}_{n,n,j} \bar{m}_n^{\text{est}})^2 - 2\mu_{n,j} (\sigma_{n,i,j} \sigma'_{n,n,j})^2 \hat{G}_{n,n,j} \bar{m}_n^{\text{est}}. \quad (57)$$

$$\frac{\mu_{i,j}(\hat{\sigma}_{i,i,j}\hat{G}_{i,i,j}\bar{m}_i^{\text{est}})^2}{\sum_{n \neq i}^N \mu_{n,j} \left(\sigma_{i,i,j} \sigma'_{i,n,j} \hat{G}_{i,i,j} \bar{m}_i^{\text{est}} \right)^2 + \frac{1}{N_t} \sum_{n=1}^N \sum_{k \neq j}^K \mu_{n,k} \hat{\sigma}_{i,i,j} \hat{G}_{i,i,j}^2 (\bar{m}'_i)^{\text{est}} B'_{n,k} + \hat{\sigma}_{i,i,j} \hat{G}_{i,i,j}^2 (\bar{m}'_i)^{\text{est}}}, \quad (58)$$

where

$$B'_{n,k} = \sigma_{i,n,k} + \hat{\sigma}_{i,i,k} (\xi_{i,k} \sigma_{i,n,k} \sigma'_{i,i,k} \hat{G}_{i,i,k} \bar{m}_i^{\text{est}})^2 - 2\mu_{i,k} (\sigma_{i,n,k} \sigma'_{i,i,k})^2 \hat{G}_{i,i,k} \bar{m}_i^{\text{est}}. \quad (59)$$

the individual terms of (63) by their asymptotic equivalents. Therefore, we assume that in the large system domain, the uplink SINR can be approximated by (58). For convenience, let us denote that denominator of (58) by $\mathcal{I}_{\text{asym}}^{\text{UL}}$. Following the same approach as in (62), we compute the uplink power allocation as the solution to the following set of equations:

$$\mu_{i,j} = \frac{\gamma_{i,j} \mathcal{I}_{\text{asym}}^{\text{UL}}}{(\hat{\sigma}_{i,i,j} \hat{G}_{i,i,j} \bar{m}_i^{\text{est}})^2} \quad \forall i, j. \quad (64)$$

The $\mu_{i,j}$ that satisfies the set of equations (64) can be computed using the following iterative method: Starting from any initial $\mu_{i,j}^0 > 0 \forall i, j$ the uplink power allocation is given by $\mu_{i,j} \triangleq \lim_{t \rightarrow \infty} \mu_{i,j}^t$, where

$$\mu_{i,j}^{t+1} = \frac{\gamma_{i,j} (\mathcal{I}_{\text{asym}}^{\text{UL}})^t}{(\hat{\sigma}_{i,i,j} \hat{G}_{i,i,j} (\bar{m}_i^{\text{est}})^t)^2} \quad \forall i, j. \quad (65)$$

In (65), $(\mathcal{I}_{\text{asym}}^{\text{UL}})^t$, $\hat{G}_{i,i,j}^t$ and $(\bar{m}_i^{\text{est}})^t$ denote the respective quantities computed at $\mu_{i,j}^t$, $\forall i, j$. We observe that numerically that the iterations of (65) converges to the solution of (64) (see numerical results in Section V).

Finally, $\bar{\delta}$ can be computed as a solution to the following linear equations

$$\bar{\delta} = \Delta^{-1} \boldsymbol{\gamma} \quad (66)$$

where the matrix $\Delta \in \mathbb{C}^{NK \times NK}$ is defined as

$$\Delta = \begin{pmatrix} \Delta^{1,1} & \dots & \Delta^{1,N} \\ \vdots & \ddots & \vdots \\ \Delta^{N,1} & \dots & \Delta^{N,N} \end{pmatrix} \quad (67)$$

where each submatrix $\Delta^{i,j} \in \mathbb{C}^{K \times K}$ is given by

$$\Delta_{j,k}^{i,n} \triangleq \begin{cases} (\hat{\sigma}_{i,i,j} \hat{G}_{i,i,j} \bar{m}_i^{\text{est}})^2, & n = i, k = j \\ -(\sigma_{n,i,j} \sigma'_{n,n,j} \hat{G}_{n,n,j} \bar{m}_n^{\text{est}})^2, & n \neq i, k = j \\ \frac{-1}{N_t} \hat{\sigma}_{n,n,k} \hat{G}_{n,n,k}^2 (\bar{m}'_n)^{\text{est}} B_{n,k}, & n = i, k \neq j \\ & \text{and } n \neq i, k \neq j. \end{cases} \quad (68)$$

V. NUMERICAL RESULTS

In this section, we present some numerical results to demonstrate the performance of the ROBF algorithm in a massive MIMO setting.

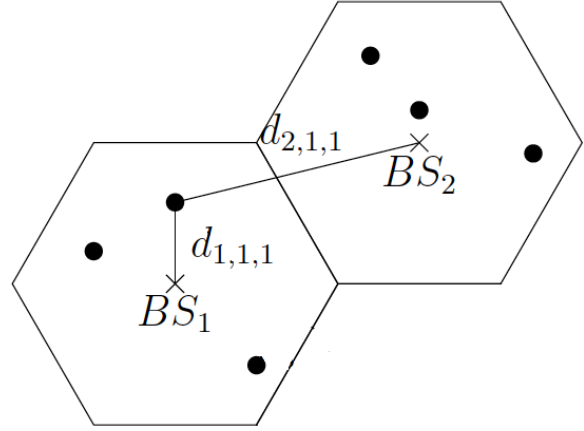


Fig. 2. Hexagonal cellular network consisting of 2 cells.

We consider a hexagonal cellular system consisting of 2 cells as shown in Figure 2. We use the distance dependent path loss model in which the path loss from BS i to UT $_{j,k}$ is given by

$$\sigma_{i,j,k} = \frac{d_0}{d_{i,j,k}^\beta},$$

where $d_{i,j,k}$ represents the distance between BS i to UT $_{j,k}$. β represents the path loss exponent which is taken to be 3.6 in all the simulation scenarios. d_0 represents the channel attenuation at reference point and is taken to $10^{-3.53}$. Location of the UTs are obtained by generating uniform random numbers inside each hexagonal cell. The distance $d_{\min} \leq d_{i,j,k} \leq d_{\max}$, where d_{\min} and d_{\max} represent the minimum and the maximum distance between the UTs to the BSs of their respective cells. In our simulation, $d_{\min} = 20$ m and $d_{\max} = \{250, 500\}$ m depending on the distance between the two BSs. The noise power is taken to be -104 dBm over the operating bandwidth. All numerical results are plotted by varying the positions of the UTs inside the cell over 500 iterations.

First, we examine the performance of the ROBF algorithm in satisfying the UT SINR constraints. Accordingly, we plot the variation of the uplink and downlink SINR (averaged across the UTs) as a function of the number of antennas per BS for 1000 channel realizations in Figure 3. Herein, $K = 50$ UTs per cell and target SINR is 3 dB per UT. The horizontal

line represents the target SINR which is 3 dB. The bubbles represent the average achieved SINR values (averaged over the channel realizations), and the vertical lines around this bubble represent the variation of the achieved downlink SINR around the average value. It can be observed that even for moderate number of antennas, e.g. 60 antennas per BS (comparable to the number of UTs), the target SINR constraints are satisfied for almost every channel realization (since the fluctuations are small). This implies that the ROBF algorithm is nearly optimal under this setting in terms of satisfying the SINR constraints.

Next, we investigate the downlink power of the ROBF algorithm and compare it with the CBF algorithm and ZF beamforming³ (denoted by $\sqrt{P_{i,j}^{ZF}} \mathbf{w}_{i,j}^{ZF}$, where $\mathbf{w}_{i,j}^{ZF}$ is unit norm vector). Note that our simulation setting, the BS performs ZF beamforming to null the interference UTs in both the cells (and not merely the UTs in its own cells). After nulling the interference, BS performs appropriate power allocation in order to meet the target SINR constraint of the UTs

$$P_{i,j}^{ZF} = \frac{\gamma_{i,j} N_0}{|\mathbf{h}_{i,i,j} \mathbf{w}_{i,j}^{ZF}|^2}.$$

We plot the downlink power per UT (i.e. $\frac{\text{Sum downlink power}}{K}$) as a function of the number of antennas per BS in Figure 4. Herein, $K = 25$ UTs per cell and target rate = 3 bits/s/Hz ($\log(1 + \gamma_{i,j})$) per UT. The following conclusions can be drawn. First, the downlink power expended by ROBF very closely matches that of CBF, indicating that the ROBF is nearly optimal in terms of minimizing the downlink power as well. Next, it can be seen that for moderate number of antennas per BS, e.g. 50 – 100 antennas, ROBF provides substantial power gains as compared to ZF beamforming. This result highlights the gains obtained by optimizing the power allocation (as compared to naively nulling out interference to all the UTs in the system). When N_t is very large, the scaling of $\frac{1}{N_t}$ starts to play a dominant role, and ROBF ceases to provide substantial gains over ZF.

Finally, we examine the efficacy of ROBF algorithm in terms of its ability to support greater number of UTs per cell. This is accomplished by fixing the number of antennas per BS to 60, and plotting the downlink power as a function of the number of UTs per cell in Figure 5. It can be seen that beyond a certain number of UTs, the downlink power corresponding to the ZF beamforming becomes unbounded. In fact, this happens at $K = 30$ UTs per cell (note that $N_t = 2K$ at this point indicating that the BS has used up all its degrees of freedom). At the same time, the downlink power of the ROBF algorithm grows unbounded at a much later stage, i.e. around 55 UTs per cell. Moreover, we can also see that as D (distance between the BSs) increases from 500 m to 1000 m, the transmit power decreases as expected. This is due to the reduced effect of inter-cell interference. Lastly, to quantify the transmit power

³Note that for eigen beamforming and RZF, after fixing the beamforming direction, the power allocation needed to satisfy the SINR constraints does not yield to a simple structure (as in the case of ZF). In order to avoid the additional complexities associated with power allocation, only ZF is used for comparison purposes. Nevertheless, we expect the gains of the ROBF algorithm to hold true with respect to the other beamforming techniques as well.

gains obtained by the ROBF algorithm over ZF, we zoom into the previous plot in Figure 6. Once again, as noted before, ROBF algorithm provides substantial power gains compared to ZF, a gain of 5 dBm per UT for $K = 25$ UTs per cell.

Impact of Pilot Contamination Effect

We finally investigate the impact of pilot contamination effect on the performance of the ROBF algorithm, and the improvements obtained by the MROBF algorithm. We consider an identical system set up as in the case of perfect CSI. Since our main objective is to characterize the performance loss due to the pilot contamination effect, we ignore the errors associated with the CSI estimation process, i.e., P_{Tr} in (47) is set to a very high value.

First, we compare the achieved SINR in the downlink by the MROBF algorithm and the CBF algorithm in Figure 7. For the implementation of CBF, we consider that Algorithm 1 is implemented with all the CSI values replaced by their estimates (i.e. $\mathbf{h}_{i,n,k}$ replaced by $\hat{\mathbf{h}}_{i,n,k} \forall i, n, k$ computed as in (47)). It can be seen that the MROBF algorithm very closely matches the target SINR values, whereas the CBF algorithm does not achieve the target SINR. This is due to the fact that in the MROBF algorithm, parameters can be computed based on the slow fading co-efficient, which can be estimated accurately at the BS. However, the CBF algorithm requires knowledge of the fast fading co-efficient, whose estimation suffers from the pilot contamination effect (in addition to the CSI estimation error that has been ignored in our numerical results).

Finally, we provide a comparison of the downlink power reduction provided by the MROBF algorithm. Notice that the CBF and ROBF do not achieve the target SINR. For the sake of comparison, we devise an algorithm that achieves the target SINR. One can think of retaining the uplink power allocation of the ROBF algorithm, and only adapting the computation of $\bar{\delta}$ as in (66), such that the target SINR is achieved in the downlink. We address this algorithm by the name ROBF with downlink adaptation. We plot the transmit power per UT as a function of the target SINR for both the MROBF algorithm and the ROBF with downlink adaptation in Figure 8. It can be seen that the transmit power per UT is significantly lower for the MROBF algorithm. Moreover, beyond a certain target SINR value, the downlink power for the ROBF with downlink adaptation becomes very high, implying that the target SINR cannot be supported by this algorithm. In contrast, the MROBF algorithm can support a greater range of target SINR values. This is due to the fact that the ROBF with downlink adaptation algorithm significantly under estimates the interference arising out of the UTs which reuse the same pilot symbols, whereas the MROBF algorithm accurately accounts for this quantity.

VI. CONCLUSION

In this work, we have formulated a decentralized multi-cell beamforming algorithm for massive MIMO systems using only locally available CSI at the BSs, and some statistical side information of the channel gains to other UTs. Our algorithm incurs a lower burden in terms of the information exchange between BSs. Further we proved theoretically that

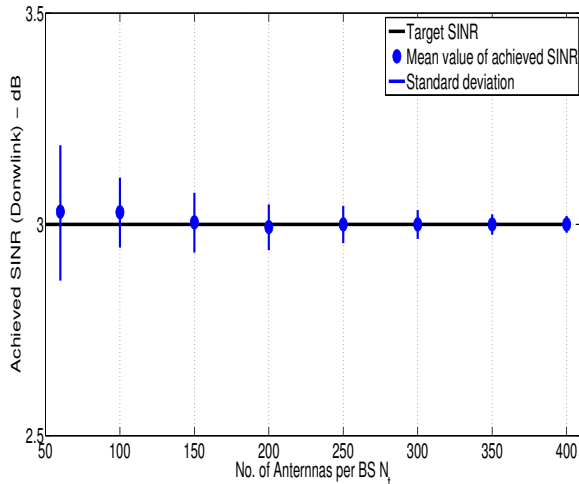


Fig. 3. Fluctuations of the downlink SINR (ROBF algorithm) around the target value. $K = 25$ UTs per cell and target SINR = 3 dB per UT.

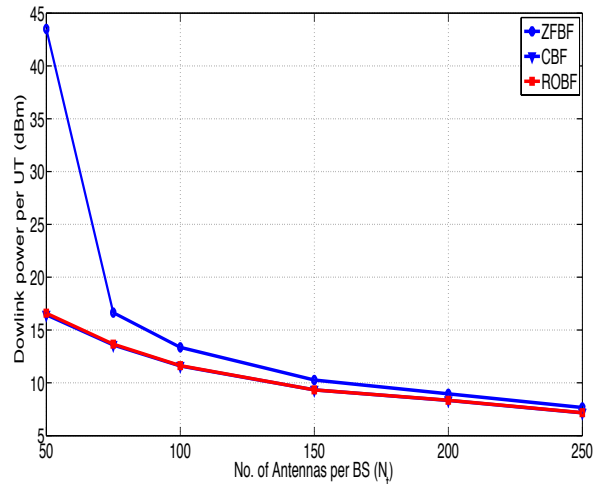


Fig. 4. Comparison of downlink power per UT as a function of the number of antennas per BS. $K = 25$ UTs per cell and target rate = 3 bits/s/Hz ($\log(1 + \gamma_{i,j})$) per UT.

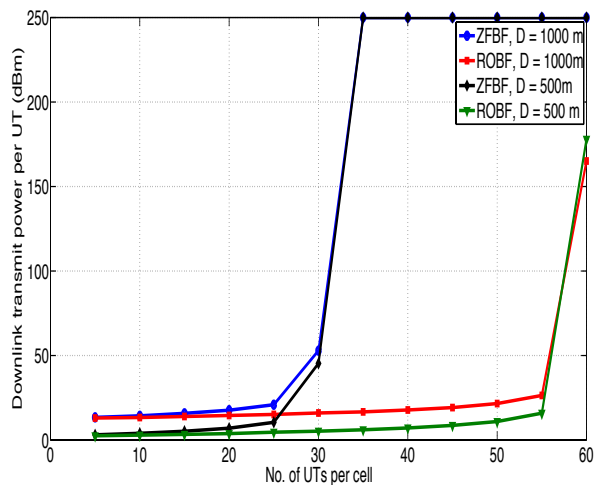


Fig. 5. Comparison of downlink power per UT as a function of the number of UTs per cell. $N_t = 60$ antennas/ BS and target rate = 3 bits/s/Hz Settings identical to Figure 5. ($\log(1 + \gamma_{i,j})$) per UT. D represents the distance between the two BSs.

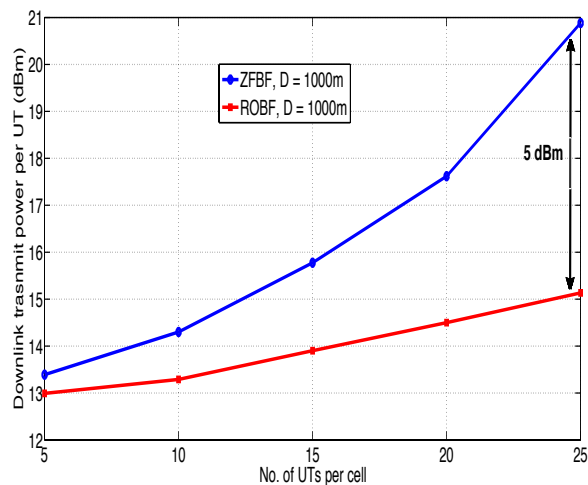


Fig. 6. Comparison of downlink power per UT as a function of the K . Settings identical to Figure 5.

this algorithm is asymptotically optimal for a large number of BS antennas and UTs. We also characterized a lower bound on the range of SINR values for which this algorithm is feasible. We confirmed using numerical results that the algorithm is nearly optimal in the regime of massive MIMO systems. It also provides substantial power savings as compared to zero-forcing beamforming when the number of antennas per BS is of the same orders of magnitude as the number of UTs per cell. Finally, we investigated the impact of CSI estimation errors and pilot contamination on the performance of the ROBF algorithm. Further, we proposed a heuristic modification of the ROBF algorithm, and showed with the help of numerical results that such an algorithm is more robust to the effects of CSI estimation errors and pilot contamination (as compared to an algorithm that uses fast fading CSI values to compute the system parameters)

VII. ACKNOWLEDGMENTS

The authors would like to thank Jakob Hoydis for his helpful discussions at various stages of this work.

This research has been supported by the ERC Starting Grant 305123 MORE.

APPENDIX A: SOME RELEVANT RESULTS

Lemma 7. ([24]) For two matrices \mathbf{X} and \mathbf{Y} , if $0 \leq \mathbf{X} \leq \mathbf{Y}$, then $\rho(\mathbf{X}) \leq \rho(\mathbf{Y})$.

Lemma 8. (Equation 2.2, [25]) Let \mathbf{A} be a Hermitian invertible matrix of size $N \times N$, then for any vector $\mathbf{x} \in \mathbb{C}^N$ and scalar $\tau \in \mathbb{C}$ for which $\mathbf{A} + \tau\mathbf{x}\mathbf{x}^H$ is invertible,

$$(\mathbf{A} + \tau\mathbf{x}\mathbf{x}^H)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\tau\mathbf{x}\mathbf{x}^H\mathbf{A}^{-1}}{1 + \tau\mathbf{x}^H\mathbf{A}^{-1}\mathbf{x}},$$

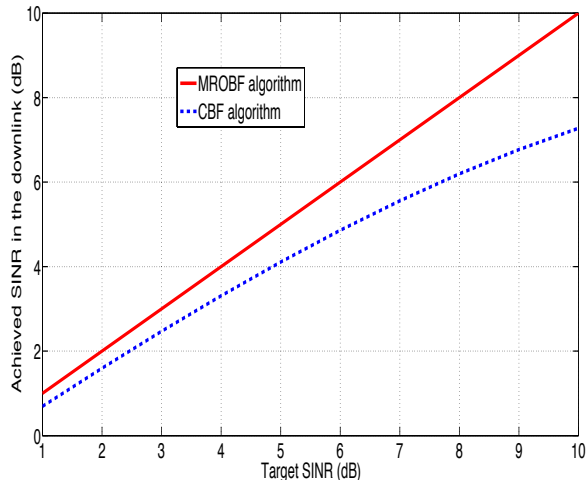


Fig. 7. Achieved SINR Vs Target SINR for MROBF and CBF algorithm in the presence of pilot contamination effect for $N_t = 60$ antennas/BS and $K = 30$ UTs per cell.

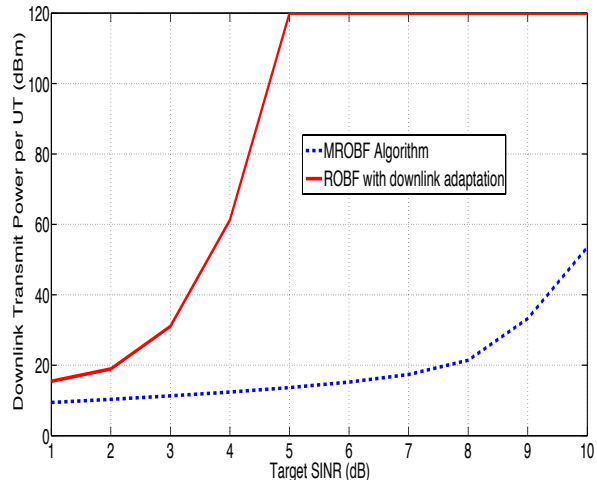


Fig. 8. Comparison of downlink power per UT Vs Target SINR for the MROBF and ROBF algorithm with downlink adaptation.

and

$$\mathbf{x}^H (\mathbf{A} + \tau \mathbf{x} \mathbf{x}^H)^{-1} = \frac{\mathbf{x}^H \mathbf{A}^{-1}}{1 + \tau \mathbf{x}^H \mathbf{A}^{-1} \mathbf{x}}.$$

Lemma 9. (Lemma 2.6, [26]) Let $\mathbf{x}, \mathbf{y} \sim \mathcal{CN}(0, \frac{1}{N} \mathbf{I}_N) \in \mathbb{C}^N$, $\mathbf{A} \in \mathbb{C}^{N \times N}$ Hermitian matrix, and that \mathbf{x} and \mathbf{y} are mutually independent, and independent of \mathbf{A} . Consider $m \geq 2$. Then there exists a constant C_m independent of N and \mathbf{A} such that

$$\mathbb{E} \left[\left| \mathbf{x}^H \mathbf{A} \mathbf{x} - \frac{1}{N} \text{tr}(\mathbf{A}) \right|^m \right] \leq \frac{C_m}{N^{m/2}} \|\mathbf{A}\|^m.$$

This implies by the Markov inequality and the Borel Cantelli lemma for $\|\mathbf{A}\| < \infty$, that

$$\mathbf{x}^H \mathbf{A} \mathbf{x} - \frac{1}{N} \text{tr}(\mathbf{A}) \xrightarrow[N \rightarrow \infty]{a.s.} 0.$$

Additionally,

$$\mathbf{x}^H \mathbf{A} \mathbf{y} \xrightarrow[N \rightarrow \infty]{a.s.} 0.$$

Lemma 10. (Lemma 2.6, [25]) Let $z \in \mathbb{C}^+$ with $v = \text{Im}(z)$ and \mathbf{A} and \mathbf{B} are $N \times N$ matrices with \mathbf{B} being Hermitian, $\tau \in \mathbb{R}$, and $\mathbf{q} \in \mathbb{C}^N$, then

$$\left| \text{tr}((\mathbf{B} - z\mathbf{I})^{-1} - (\mathbf{B} + \tau \mathbf{q} \mathbf{q}^H - z\mathbf{I})^{-1}) \mathbf{A} \right| \leq \frac{\|\mathbf{A}\|}{v}.$$

Theorem 5. ([25]) Consider the matrix $\mathbf{B}_{N_t} = \mathbf{X}_{N_t} \mathbf{T}_{N_t} \mathbf{X}_{N_t}^H$, where $\mathbf{X}_{N_t} = \frac{1}{\sqrt{N_t}} \mathbf{Y}^{N_t} \in \mathbb{C}^{N_t \times NK}$ with entries $\mathbf{Y}^{N_t}(p, q) \sim \mathcal{CN}(0, 1)$, and the matrix \mathbf{T}_{N_t} a non random diagonal matrix given by $\mathbf{T}_{N_t} = \text{diag}(t_1, \dots, t_{NK}) \in \mathbb{R}^{NK \times NK}$. Let $m(z) = \frac{1}{N_t} \text{tr}(\mathbf{B}_{N_t} + z\mathbf{I})^{-1}$, $z > 0$, Then, $m(z) - \bar{m}(z) \xrightarrow[N_t, K \rightarrow \infty]{a.s.} 0$, where $\bar{m}(z)$ can be evaluated as the unique solution to the fixed point equation

$$\bar{m}(z) = \left(\frac{1}{N_t} \sum_{i=1}^{NK} \frac{t_i}{1 + t_i \bar{m}(z)} + z \right)^{-1}.$$

Further from [14], under the same assumptions as Theorem 5, let $m'(z) = \frac{1}{N_t} \text{tr}(\mathbf{B}_{N_t} + z\mathbf{I})^{-2}$. Then,

$$m'(z) - \bar{m}'(z) \xrightarrow[N_t, K \rightarrow \infty]{a.s.} 0 \quad (69)$$

where $\bar{m}'(z)$ can be evaluated as

$$\bar{m}'(z) = \frac{\bar{m}^2(z)}{1 - \frac{\bar{m}^2(z)}{N_t} \sum_{i=1}^{NK} \frac{t_i^2}{(1 + t_i \bar{m}(z))^2}}. \quad (70)$$

At this point, we remark that all the above results also hold for a more general non-Gaussian vectors/ matrices satisfying some moment conditions.

Lemma 11. (Lemma 1, [27]) Let a_N, b_N, x_N and y_N denote four infinite sequences of complex random variables indexed by N . If $a_N \asymp b_N$ and $x_N \asymp y_N$ and if $|a_N|, |y_N|^{-1}$ and/or $|a_N|, |y_N|^{-1}$ are uniformly bounded above over N , then $a_N/x_N \asymp b_N/y_N$. The requirement of uniform bound can be relaxed to boundedness almost surely.

APPENDIX B: PROOF OF THE FIXED POINT EQUATION

In order to prove the convergence of the iterative equation (11), we use the arguments of standard function [28]. A K -variate function $\mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}), \dots, g_K(\mathbf{x})] \in \mathbb{R}^K$ for $\mathbf{x} \in \mathbb{C}^K$ is said to be standard if it fulfills the following conditions:

- Positivity: $\mathbf{g}(\mathbf{x}) > 0$, for $\mathbf{x} \geq 0$.
- Scalability: For $\beta > 1$, $\beta \mathbf{g}(\mathbf{x}) > \mathbf{g}(\beta \mathbf{x})$.
- Monotonicity: For $\mathbf{x}' \geq \mathbf{x}$, $\mathbf{g}(\mathbf{x}') \geq \mathbf{g}(\mathbf{x})$.

Now consider the following iterative algorithm given by,

$$\mathbf{x}^{t+1} = \mathbf{g}(\mathbf{x}^t), \quad t \geq 1 \quad (71)$$

If the K -variate function is standard, then it ensures the convergence of (71) to its unique fixed point solution $\mathbf{x} = \mathbf{g}(\mathbf{x})$, if the solution exists.

Consider the iterative equations in (11) which can be represented as

$$\mu_{i,j}^{t+1} = f_{i,j}(\boldsymbol{\mu}^t) \quad (72)$$

where $\boldsymbol{\mu}_i^t = [\mu_{i,1}^t, \dots, \mu_{i,K}^t]^T$ and $\boldsymbol{\mu}^t = [\boldsymbol{\mu}_1^t, \dots, \boldsymbol{\mu}_N^t]^T$ and

$$f_{i,j}(\boldsymbol{\mu}^t) \triangleq \frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i^t} \quad \forall i, j. \quad (73)$$

Let us also define the NK -variate function

$$\begin{aligned} \mathbf{f}_i(\boldsymbol{\mu}^t) &\triangleq [f_{i,1}(\boldsymbol{\mu}^t), \dots, f_{i,K}(\boldsymbol{\mu}^t)]^T \in \mathbb{R}^K \\ \mathbf{f}(\boldsymbol{\mu}^t) &\triangleq [\mathbf{f}_1(\boldsymbol{\mu}^t), \dots, \mathbf{f}_N(\boldsymbol{\mu}^t)]^T \in \mathbb{R}^{NK}. \end{aligned} \quad (74)$$

The existence of the fixed point to the equation $\boldsymbol{\mu} = \mathbf{f}(\boldsymbol{\mu})$ follows from Lemma 1. We will now prove that the NK -variate function $\mathbf{f}(\boldsymbol{\mu})$ is a standard function and hence the iterations of (72) converges to its unique fixed point. In the subsequent part of this proof, we will introduce the notation $\bar{m}_i^t(z)$ to denote the solution of the fixed point equation

$$\bar{m}_i^t(z) = \left(\frac{1}{N_t} \sum_{n=1}^N \sum_{k=1}^K \frac{\sigma_{i,n,k} \mu_{n,k}^t}{1 + \sigma_{i,n,k} \mu_{n,k}^t \bar{m}_i^t(z)} + z \right)^{-1}. \quad (75)$$

With this notation, the solution to the fixed point equation of (12) can be written as $\bar{m}_i^t(1)$. For notational convenience, we will also drop the superscript t .

Positivity: The positivity result follows directly since $\bar{m}_i(1)$ is positive whenever $\boldsymbol{\mu} \geq 0$. Hence $f_{i,j}(\boldsymbol{\mu}) > 0$.

Scalability: Let us consider the difference between the following quantities.

$$\begin{aligned} \beta f_{i,j}(\boldsymbol{\mu}) - f_{i,j}(\beta \boldsymbol{\mu}) &= \frac{\beta \gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i^{(1)}(1)} - \frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i^{(2)}(1)} \\ &= \frac{\gamma_{i,j}}{\sigma_{i,i,j}} \left(\frac{\beta}{\bar{m}_i^{(1)}(1)} - \frac{1}{\bar{m}_i^{(2)}(1)} \right) \end{aligned} \quad (76)$$

where $\beta > 1$ and $\bar{m}_i^{(1)}(1)$ and $\bar{m}_i^{(2)}(1)$ are the unique solutions to (12) evaluated at $\boldsymbol{\mu}$ and $\beta \boldsymbol{\mu}$. In order to evaluate $\bar{m}_i^{(2)}(1)$, we go back to the definition $\bar{m}_i^{(2)}(1)$ evaluated at $\beta \boldsymbol{\mu}$.

$$\bar{m}_i^{(2)}(1) = \frac{1}{N} \text{tr}(\mathbf{H}_i \beta \mathbf{M} \mathbf{H}_i^H + \mathbf{I})^{-1} \asymp \frac{1}{\beta} \bar{m}_i \left(\frac{1}{\beta} \right) \quad (77)$$

where $\mathbf{M} = \text{diag}[\boldsymbol{\mu}]$. Clearly $1/\beta < 1$. $\bar{m}_i^{(2)}(1)$ can be evaluated as the solution to the fixed point equation (12) evaluated at the point $1/\beta$ and then scaling the result by β . From (76) and (77), it can be concluded that in order to prove the scalability result, it is sufficient to show that $\bar{m}_i(z)$ is a decreasing function of z .

In order to prove the same, let us consider an extended version of the channel matrix which is constructed as follows. Defining, $\mathbf{R}_{i,j,k} = \sigma_{i,j,k} \mathbf{I}_{N_t} \in \mathbb{R}^{N_t \times N_t}$ and $\mathbf{R}_{i,j,k}^L = \sigma_{i,j,k} \mathbf{I}_{N_t L} \in \mathbb{R}^{N_t L \times N_t L}$. The matrix \mathbf{H}_i^L is constructed as follows:

$$\begin{aligned} \mathbf{H}_{i,j}^L &= \frac{1}{\sqrt{L}} \left[\mathbf{R}_{i,j,1}^{1/2} \mathbf{X}_{i,j}^L, \dots, \mathbf{R}_{i,j,K}^{1/2} \mathbf{X}_{i,j}^L \right] \in \mathbb{C}^{N_t L \times K L} \\ \mathbf{H}_i^L &= [\mathbf{H}_{i,1}^L, \dots, \mathbf{H}_{i,N}^L] \in \mathbb{C}^{N_t L \times N K L} \end{aligned} \quad (78)$$

where the matrix $\mathbf{X}_{i,j}^L \in \mathbb{C}^{N_t \times K}$, whose elements $\mathbf{X}_{i,j}^L(p, q) \sim \mathcal{CN}(0, \frac{1}{N_t})$. Also, let us define the following,

$$\begin{aligned} \boldsymbol{\lambda}_{i,j}^L &= [\lambda_{i,j}, \dots, \lambda_{i,j}]^T \in \mathbb{R}^{L \times 1} \\ \boldsymbol{\lambda}_i^L &= [\boldsymbol{\lambda}_{i,1}^L, \dots, \boldsymbol{\lambda}_{i,K}^L]^T \in \mathbb{R}^{K L \times 1} \\ \boldsymbol{\Lambda}^L &= \text{diag}(\boldsymbol{\lambda}_1^L, \dots, \boldsymbol{\lambda}_N^L) \in \mathbb{R}^{N K L \times N K L} \end{aligned} \quad (79)$$

and let $\mathbf{Q}_i(z) = (\mathbf{H}_i \boldsymbol{\Lambda} \mathbf{H}_i^H + z \mathbf{I}_{N_t L})^{-1}$ and $\mathbf{Q}_i^L(z) = (\mathbf{H}_i^L \boldsymbol{\Lambda}^L (\mathbf{H}_i^L)^H + z \mathbf{I}_{N_t L})^{-1}$. Let us denote

$$m_i^L(z) = \frac{1}{N_t L} \text{tr}(\mathbf{Q}_i^L(z)). \quad (80)$$

It can be verified that for any fixed N, N_t and K , the following limit holds,

$$m_i^L(z) - \bar{m}_i(z) \xrightarrow[L \rightarrow \infty]{\text{a.s.}} 0. \quad (81)$$

Now consider the difference between the following two quantities, for any $z_2 > z_1 > 0$ and for any positive L , we have

$$\begin{aligned} m_i^L(z_1) - m_i^L(z_2) &= \frac{1}{N_t L} \text{tr}(\mathbf{Q}_i^L(z_1) - \mathbf{Q}_i^L(z_2)) \\ &\stackrel{(a)}{>} 0 \end{aligned} \quad (82)$$

where inequality (a) follows by the following identity: For invertible matrices \mathbf{A} and \mathbf{B} ,

$$\mathbf{A}^{-1} - \mathbf{B}^{-1} = -\mathbf{A}^{-1}(\mathbf{A} - \mathbf{B})\mathbf{B}^{-1}. \quad (83)$$

We will now show that for any fixed N, N_t and K , the inequality in (82) of the random quantities ($m_i^L(z)$) also hold for their respective deterministic approximations ($\bar{m}_i(z)$). This can be argued as follows. Consider the difference

$$\begin{aligned} m_i^L(z_1) - m_i^L(z_2) &= m_i^L(z_1) - \bar{m}_i(z_1) - m_i^L(z_2) \\ &\quad + \bar{m}_i(z_2) + \bar{m}_i(z_1) - \bar{m}_i(z_2). \end{aligned} \quad (84)$$

First note that since the matrix \mathbf{H}_i^L has bounded spectral norm almost surely, it follows that, almost surely,

$$\lim_{L \rightarrow \infty} m_i^L(z) > 0. \quad (85)$$

From (85), further applying the result of (81) in the right hand side of (84) and the inequality of (82), it follows that

$$\bar{m}_i(z_1) - \bar{m}_i(z_2) > 0. \quad (86)$$

Thus, $\bar{m}_i(z)$ is a decreasing function of z . We remark that the strict positivity of the $m_i^L(z)$ in (85) is essential for the strict positivity argument of (86).

Monotonicity: Consider $\boldsymbol{\mu}' \geq \boldsymbol{\mu}$. In this case, we denote $\bar{m}_i^{(1)}(1)$ and $\bar{m}_i^{(2)}(1)$ as the solutions to the fixed point equations in (12) evaluated at $\boldsymbol{\mu}'$ and $\boldsymbol{\mu}$ respectively. As before, let us consider the difference between the quantities,

$$f_{i,j}(\boldsymbol{\mu}') - f_{i,j}(\boldsymbol{\mu}) = \frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i^{(1)}(1)} - \frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i^{(2)}(1)}. \quad (87)$$

We now have to show that $\bar{m}_i^{(1)}(1) \leq \bar{m}_i^{(2)}(1)$ in order to prove the monotonicity result. This can be shown by

constructing the extended matrices as in (78) and noting the monotonicity property for the associated random quantities and extending the result to their deterministic approximations for any system dimensions. The proof is similar to the scalability result and hence omitted here.

APPENDIX C: FEASIBILITY CONDITIONS

A. Part I: Proof of Lemma 2

Let us first consider the expression for the achieved uplink SINR by the ROBF algorithm given by

$$\Lambda_{i,j}(\boldsymbol{\mu}) = \frac{\frac{\mu_{i,j}}{N_t} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,i,j}|^2}{\sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k}}{N_t} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,n,k}|^2 + \|\hat{\mathbf{v}}_{i,j}\|_2^2}. \quad (88)$$

We examine the expression for the uplink SINR in the large system regime. It can be shown that the expression for the uplink SINR converges asymptotically to the following (the proof of this is very similar to the convergence result of the downlink interference and downlink SINR provided in Lemma 4 and Appendix E, and hence omitted in order to avoid repetition):

$$\begin{aligned} \Lambda_{i,j}(\boldsymbol{\mu}) &\asymp \frac{\mu_{i,j} \sigma_{i,i,j} \bar{G}_{i,i,j} \bar{m}_i^2}{\sum_{(n,k) \neq (i,j)} \frac{1}{N_t} \mu_{n,k} \bar{G}_{i,n,k} \bar{G}_{i,i,j} \bar{m}_i' + \alpha_i \bar{G}_{i,i,j} \bar{m}_i'} \\ &= \frac{\mu_{i,j} \sigma_{i,i,j} \bar{m}_i^2}{\sum_{(n,k) \neq (i,j)} \frac{1}{N_t} \mu_{n,k} \bar{G}_{i,n,k} \bar{m}_i' + \alpha_i \bar{m}_i'}. \end{aligned} \quad (89)$$

Let us we consider a slightly modified version of the expression (89) as follows:

$$\frac{\mu_{i,j} \sigma_{i,i,j} \bar{m}_i^2}{\sum_{n,k} \frac{1}{N_t} \mu_{n,k} \bar{G}_{i,n,k} \bar{m}_i' + \alpha_i \bar{m}_i'}. \quad (90)$$

We now examine the expression (90) in detail. Recall the fixed point equation for the computation of \bar{m}_i in (12). Upon rearranging the terms, we have

$$\bar{m}_i = 1 - \frac{1}{N_t} \sum_{n,k} \frac{\sigma_{i,n,k} \mu_{n,k} \bar{m}_i}{1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i}. \quad (91)$$

Substituting the expressions for $\bar{G}_{i,n,k}$, \bar{m}_i and \bar{m}_i' in (90), we get

$$(90) = \frac{\sigma_{i,i,j} \mu_{i,j} \left(1 - \frac{1}{N_t} \sum_{n,k} \frac{(\sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}\right)}{1 + \frac{1}{N_t} \sum_{n,k} \frac{\sigma_{i,n,k} \mu_{n,k}}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}}. \quad (92)$$

Multiplying and diving by \bar{m}_i in (92), we obtain

$$\begin{aligned} (90) &= \frac{\sigma_{i,i,j} \mu_{i,j} \bar{m}_i \left(1 - \frac{1}{N_t} \sum_{n,k} \frac{(\sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}\right)}{\bar{m}_i + \frac{1}{N_t} \sum_{n,k} \frac{\sigma_{i,n,k} \mu_{n,k} \bar{m}_i}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}} \\ &= \frac{\sigma_{i,i,j} \mu_{i,j} \bar{m}_i \left(1 - \frac{1}{N_t} \sum_{n,k} \frac{(\sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}\right)}{1 - \frac{1}{N_t} \sum_{n,k} \frac{\sigma_{i,n,k} \mu_{n,k} \bar{m}_i}{1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i} + \frac{1}{N_t} \sum_{n,k} \frac{\sigma_{i,n,k} \mu_{n,k} \bar{m}_i}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}} \\ &= \frac{\sigma_{i,i,j} \mu_{i,j} \bar{m}_i \left(1 - \frac{1}{N_t} \sum_{n,k} \frac{(\sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}\right)}{1 - \frac{1}{N_t} \sum_{n,k} \frac{(\sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}} \\ &= \sigma_{i,i,j} \mu_{i,j} \bar{m}_i. \end{aligned} \quad (93)$$

We notice that that the expression at (93) is exactly in the same form as the right hand side of the fixed point equation (19). Therefore, we can rewrite (19) as follows:

$$\gamma_{i,j} = \frac{\mu_{i,j} \sigma_{i,i,j} \bar{G}_{i,i,j} \bar{m}_i^2}{\sum_{n,k} \frac{1}{N_t} \mu_{n,k} \bar{G}_{i,n,k} \bar{G}_{i,i,j} \bar{m}_i' + \bar{G}_{i,i,j} \bar{m}_i'} \quad \forall i, j.$$

Rearranging, we have

$$\begin{aligned} \frac{1}{\gamma_{i,j}} \mu_{i,j} \sigma_{i,i,j} \bar{G}_{i,i,j} \bar{m}_i^2 &= \sum_{n,k} \frac{\mu_{n,k}}{N_t} \bar{G}_{i,n,k} \bar{G}_{i,i,j} \bar{m}_i' \\ &\quad + \bar{G}_{i,i,j} \bar{m}_i' \quad \forall i, j \\ \Rightarrow \mu_{i,j} &= \sum_{n,k} \frac{\gamma_{i,j} \mu_{n,k}}{N_t} \frac{\bar{G}_{i,n,k} \bar{G}_{i,i,j} \bar{m}_i'}{\sigma_{i,i,j} \bar{G}_{i,i,j} \bar{m}_i^2} \\ &\quad + \frac{\bar{G}_{i,i,j} \bar{m}_i'}{\sigma_{i,i,j} \bar{G}_{i,i,j} \bar{m}_i^2} \quad \forall i, j. \end{aligned}$$

The equation in matrix form can be written as

$$\boldsymbol{\mu} = \boldsymbol{\Gamma}(\boldsymbol{\Delta}')^T \boldsymbol{\mu} + \boldsymbol{\kappa} \quad (94)$$

where $\boldsymbol{\kappa}_i = \left[\frac{\bar{m}_i'}{\sigma_{i,i,1} \bar{m}_i^2}, \dots, \frac{\bar{m}_i'}{\sigma_{i,i,K} \bar{m}_i^2} \right]^T$ and $\boldsymbol{\kappa} = [\kappa_1, \dots, \kappa_N]^T$. The matrix $\boldsymbol{\Gamma}$ is the same as the matrix $\boldsymbol{\Gamma}$ defined in the description of the ROBF algorithm. The matrix $\boldsymbol{\Delta}'$ is defined as follows.

$$\boldsymbol{\Delta}' = \begin{pmatrix} (\boldsymbol{\Delta}')^{1,1} & \dots & (\boldsymbol{\Delta}')^{1,N} \\ \vdots & \ddots & \vdots \\ (\boldsymbol{\Delta}')^{N,1} & \dots & (\boldsymbol{\Delta}')^{N,N} \end{pmatrix} \quad (95)$$

where each submatrix $(\boldsymbol{\Delta}')^{i,j} \in \mathbb{C}^{K \times K}$ is given by

$$(\boldsymbol{\Delta}')_{j,k}^{i,n} \triangleq \begin{cases} \frac{1}{N_t} \bar{G}_{i,i,j} \bar{G}_{i,i,j} \bar{m}_i', & n = i, k = j \\ \frac{1}{N_t} \bar{G}_{i,i,j} \bar{G}_{i,i,k} \bar{m}_i', & n = i, k \neq j \\ \frac{1}{N_t} \bar{G}_{n,i,j} \bar{G}_{n,n,k} \bar{m}_n', & n \neq i. \end{cases} \quad (96)$$

Notice that if the SINR targets are feasible for the uplink solution, the linear equations in (94) must have a solution. In other words, the matrix $\mathbf{I} - \boldsymbol{\Gamma}(\boldsymbol{\Delta}')^T$ must be invertible. Problems with similar structure has been studied before in the context of power allocation in wireless networks [29]. It has been established that in such problems that invertibility of the matrix is ensured if $\rho(\boldsymbol{\Gamma}(\boldsymbol{\Delta}')^T) < 1$ (where ρ is the spectral radius of the matrix) [29]. Now recall the matrix $\boldsymbol{\Delta}$ defined in equation (16). Observe that the matrix $\boldsymbol{\Delta}$ and $\boldsymbol{\Delta}'$ only differ in the diagonal element. Also, we can note that $\boldsymbol{\Delta} \leq \boldsymbol{\Delta}'$ (denotes element wise inequality, refer to the notations), and hence $\boldsymbol{\Gamma} \boldsymbol{\Delta} \leq \boldsymbol{\Gamma} \boldsymbol{\Delta}'$. Hence, it follows from Lemma 7, that $\rho(\boldsymbol{\Gamma}(\boldsymbol{\Delta})^T) \leq \rho(\boldsymbol{\Gamma}(\boldsymbol{\Delta}')^T) < 1$.

Now recall the linear equations for computing $\bar{\boldsymbol{\delta}}$ given by

$$\bar{\boldsymbol{\delta}} = \boldsymbol{\Gamma} \boldsymbol{\Delta} \bar{\boldsymbol{\delta}} + \boldsymbol{\rho}.$$

It has been established in works before [30], the eigen values of the matrices $\boldsymbol{\Gamma} \boldsymbol{\Delta}$ and $\boldsymbol{\Gamma} \boldsymbol{\Delta}'$ are the same. For completeness, this can be argued as follows,

$$\begin{aligned} |\boldsymbol{\Gamma} \boldsymbol{\Delta} - \lambda \mathbf{I}| &= |\boldsymbol{\Gamma}| |\boldsymbol{\Delta} - \lambda \boldsymbol{\Gamma}^{-1}| = |\boldsymbol{\Gamma}| |(\boldsymbol{\Delta} - \lambda \boldsymbol{\Gamma}^{-1})^T| \\ &= |\boldsymbol{\Gamma}| |\boldsymbol{\Delta}^T - \lambda \boldsymbol{\Gamma}^{-1}| = |\boldsymbol{\Gamma} \boldsymbol{\Delta}^T - \lambda \mathbf{I}| = 0. \end{aligned}$$

From the above discussion, we conclude that the spectral radius of $\Gamma\Delta$ and $\Gamma\Delta^T$ are the same and, hence, if $\rho(\Gamma\Delta) < 1$ then $\rho(\Gamma\Delta^T) < 1$ and vice versa. Consequently, the matrix $I - \Gamma\Delta$ is invertible. This proves the result of Lemma 2.

B. Part II: Feasible SINR targets for the modified system

Consider the equation (25). Rearranging (25), we obtain

$$\begin{aligned} 1 &= \frac{1}{N_t} \sum_{k=1}^K \frac{\sigma_{i,i,k} \mu_{i,k}^{\text{mod}} \bar{m}_i^{\text{mod}}}{1 + \sigma_{i,i,k} \mu_{i,k}^{\text{mod}} \bar{m}_i^{\text{mod}}} \\ &\quad + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\sigma_{\max}(n) \mu_{n,k}^{\text{mod}}}{1 + \sigma_{\max}(n) \mu_{n,k}^{\text{mod}} \bar{m}_i^{\text{mod}}} + \bar{m}_i^{\text{mod}} \\ &= \frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} \\ &\quad + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\sigma_{\max}(n) \mu_{n,k}^{\text{mod}} \bar{m}_i^{\text{mod}}}{1 + \sigma_{\max}(n) \mu_{n,k}^{\text{mod}} \bar{m}_i^{\text{mod}}} + \bar{m}_i^{\text{mod}}. \end{aligned} \quad (97)$$

Equation (97) is true for all $\bar{m}_i^{\text{mod}}, i = 1, \dots, N$. Therefore, by symmetry of the fixed point equation $\bar{m}_i^{\text{mod}} = \bar{m}^{\text{mod}}, i = 1, \dots, N$. Thus,

$$\sigma_{\max}(n) \mu_{n,k}^{\text{mod}} \bar{m}_i^{\text{mod}} = \sigma_{\max}(n) \mu_{n,k}^{\text{mod}} \bar{m}^{\text{mod}} = \frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}. \quad (98)$$

Substituting (98) in (97) yields

$$1 = \frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}{1 + \frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}} + \bar{m}_i. \quad (99)$$

Now recall (24). Rearranging, we obtain

$$\mu_{i,j}^{\text{mod}} = \frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i^{\text{mod}}}. \quad (100)$$

The feasibility condition **[C1]** requires that $\mu_{i,j}^{\text{mod}} < \infty, \forall i, j$. This implies from (100) that $\bar{m}_i^{\text{mod}} > 0$. Using this in (99), we obtain the following condition on the target SINR for feasibility:

$$\frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}{1 + \frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}} < 1. \quad (101)$$

1) *Proof of Corollary 1:* Let us rewrite (99) as

$$\bar{m}_i = 1 - \frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}{1 + \frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}. \quad (102)$$

The feasibility of the uplink problem in the asymptotic domain implies that $\limsup_{N_t, K \rightarrow \infty} \mu_{i,j}^{\text{mod}} < \infty$. This in turn implies that $\liminf_{N_t, K \rightarrow \infty} \bar{m}_i^{\text{mod}} > 0$ (strictly positive).

Thus, we consider (102) in the asymptotic regime, and take $\liminf_{N_t, K \rightarrow \infty}$. This yields,

$$\begin{aligned} \liminf_{N_t, K \rightarrow \infty} \bar{m}_i &= \liminf_{N_t, K \rightarrow \infty} \left[1 - \frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} \right. \\ &\quad \left. - \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}{1 + \frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}} \right]. \end{aligned} \quad (103)$$

$$\begin{aligned} &= 1 - \limsup_{N_t, K \rightarrow \infty} \left[\frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} \right. \\ &\quad \left. - \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}{1 + \frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}} \right] \end{aligned} \quad (104)$$

Since $\liminf_{N_t, K \rightarrow \infty} \bar{m}_i^{\text{mod}} > 0$, and (104), it follows that the feasibility conditions are given by,

$$\begin{aligned} \limsup_{N_t, K \rightarrow \infty} \left[\frac{1}{N_t} \sum_{k=1}^K \frac{\gamma_{i,k}}{1 + \gamma_{i,k}} + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}}{1 + \frac{\sigma_{\max}(n)}{\sigma_{n,n,k}} \gamma_{n,k}} \right] \\ < 1 \quad \forall i. \end{aligned} \quad (105)$$

C. Part III: Proof of Lemma 3

Let us denote $\boldsymbol{\mu}_i^{\text{mod}} = [\mu_{i,1}^{\text{mod}}, \dots, \mu_{i,K}^{\text{mod}}]^T$ and $\boldsymbol{\mu}^{\text{mod}} = [\boldsymbol{\mu}_1^{\text{mod}}, \dots, \boldsymbol{\mu}_N^{\text{mod}}]^T$. The main idea behind this proof is the following: Consider the pair of vector $\{\boldsymbol{\gamma}, \boldsymbol{\mu}^{\text{mod}}\}$ that satisfies (24). We show that when the power allocation $\boldsymbol{\mu}^{\text{mod}}$ is used in the original system, then

$$\sigma_{i,i,j} \mu_{i,j}^{\text{mod}} \bar{m}_i^{\text{mod}1} \geq \gamma_{i,j} \quad \forall i, j \quad (106)$$

where $\bar{m}_i^{\text{mod}1}$ satisfies

$$\begin{aligned} \frac{1}{\bar{m}_i^{\text{mod}1}} &= \frac{1}{N_t} \frac{\sigma_{i,i,k} \mu_{i,k}^{\text{mod}}}{1 + \sigma_{i,i,k} \mu_{i,k}^{\text{mod}} \bar{m}_i^{\text{mod}1}} \\ &\quad + \frac{1}{N_t} \sum_{\substack{n=1 \\ n \neq i}}^N \sum_{k=1}^K \frac{\sigma_{i,n,k} \mu_{n,k}^{\text{mod}}}{1 + \sigma_{i,n,k} \mu_{n,k}^{\text{mod}} \bar{m}_i^{\text{mod}1}} + 1, \end{aligned} \quad (107)$$

where (107) represents the fixed point equation calculated by utilizing the power allocation $\boldsymbol{\mu}^{\text{mod}}$ in the original system. If (106) is true, then this implies that $\boldsymbol{\gamma}$ is a feasible SINR target vector for the original system (since there exists a power allocation $\boldsymbol{\mu}^{\text{mod}}$ that can achieve this SINR target).

We prove this as follows: Firstly, it is easy to see that

$$\sigma_{\max}(n) \mu_{n,k}^{\text{mod}} \geq \sigma_{i,n,k} \mu_{n,k}^{\text{mod}} \quad \forall n, k \quad (108)$$

since $\sigma_{\max}(n) \geq \sigma_{i,n,k}$. Recall that the function $f_{i,j} \triangleq \frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i}$ is a standard function. From (108) and the monotonicity property of standard functions, we have

$$\frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i^{\text{mod}}} \geq \frac{\gamma_{i,j}}{\sigma_{i,i,j} \bar{m}_i^{\text{mod}1}}. \quad (109)$$

Rearranging (109) we have, $\sigma_{i,i,j} \bar{m}_i^{\text{mod}1} \geq \sigma_{i,i,j} \bar{m}_i^{\text{mod}}$. Multiplying by $\mu_{i,j}^{\text{mod}}$, it follows that

$$\sigma_{i,i,j} \mu_{i,j}^{\text{mod}} \bar{m}_i^{\text{mod}1} \geq \sigma_{i,i,j} \mu_{i,j}^{\text{mod}} \bar{m}_i^{\text{mod}}. \quad (110)$$

From (24), $\sigma_{i,i,j}\mu_{i,j}^{\text{mod}}\bar{m}_i^{\text{mod}} = \gamma_{i,j}$. Therefore, we can conclude that

$$\sigma_{i,i,j}\mu_{i,j}^{\text{mod}}\bar{m}_i^{\text{mod}} \geq \gamma_{i,j}.$$

APPENDIX D: CONVERGENCE OF THE UPLINK SINR

In this Appendix, we prove the result of Lemma 3, statement (36). Recall that the expression of uplink SINR for the ROBF algorithm is given by (32). We make a variable change and denote $\mathbf{x}_{i,i,j} = \frac{\mathbf{h}_{i,i,j}}{\sqrt{N_t}}$. This implies that

$$\Lambda_{i,j}(\boldsymbol{\mu}) = \mu_{i,j}\mathbf{x}_{i,i,j}^H(\boldsymbol{\Sigma}_i^{\mu} + \mathbf{I}_{N_t})^{-1}\mathbf{x}_{i,i,j} \quad \forall i,j. \quad (111)$$

Note that $\mu_{i,j}$ is independent of the elements of the channel matrix. Applying Lemma 9 to the quadratic term of $\mathbf{x}_{i,i,j}^H(\boldsymbol{\Sigma}_i^{\mu} + \mathbf{I}_{N_t})^{-1}\mathbf{x}_{i,i,j}$ yields to,

$$\begin{aligned} \mathbb{E} \left[\left| \mathbf{x}_{i,i,j}^H(\boldsymbol{\Sigma}_i^{\mu} + \mathbf{I}_{N_t})^{-1}\mathbf{x}_{i,i,j} - \frac{1}{N_t}\text{tr}(\boldsymbol{\Sigma}_i^{\mu} + \mathbf{I}_{N_t})^{-1} \right|^k \right] \\ \leq \frac{C_1}{N_t^{\frac{k}{2}}} \quad \forall i,j \end{aligned} \quad (112)$$

for $k \geq 2$, and constant C_1 independent of N_t and K . Additionally, from the result of Lemma 6.1, [31], we have

$$\mathbb{E} \left[\left| \frac{1}{N_t}\text{tr}(\boldsymbol{\Sigma}_i^{\mu} + \mathbf{I}_{N_t})^{-1} - \bar{m}_i \right|^k \right] \leq \frac{C_2}{N_t^{\frac{k}{2}}}, \quad \forall i,j \quad (113)$$

for $k \geq 2$ constant C_2 independent of N_t and K . Therefore, from (112), (113) and Holder's inequality ($|x+y|^k \leq 2^{k-1}(|x|^k + |y|^k)$) we conclude that for some constant C_3 ,

$$\mathbb{E} \left[\left| \mathbf{x}_{i,i,j}^H(\boldsymbol{\Sigma}_i^{\mu} + \mathbf{I}_{N_t})^{-1}\mathbf{x}_{i,i,j} - \bar{m}_i \right|^k \right] \leq \frac{C_3}{N_t^{\frac{k}{2}}}, \quad \forall i,j. \quad (114)$$

From (111), we have, $\mathbf{x}_{i,i,j}^H(\boldsymbol{\Sigma}_i^{\mu} + \mathbf{I}_{N_t})^{-1}\mathbf{x}_{i,i,j} = \frac{\Lambda_{i,j}(\mu_{i,j})}{\sigma_{i,i,j}\mu_{i,j}^{\text{mod}}}$. Moreover, from the convergence of the fixed point equation (11), $\bar{m}_i = \frac{\gamma_{i,j}}{\sigma_{i,i,j}\mu_{i,j}^{\text{mod}}}$. Substituting in (114) we have,

$$\mathbb{E} \left[\left| \Lambda_{i,j}(\mu_{i,j}) - \gamma_{i,j} \right|^k \right] \leq \frac{C_4}{N_t^{\frac{k}{2}}}, \quad \forall i,j \quad (115)$$

where $C_4 = C_3\sigma_{i,i,j}\mu_{i,j}$. Note that $\sigma_{i,i,j}$ is bounded. Moreover for any target SINR γ satisfying the feasibility conditions of (27), $\limsup_{N_t, K \rightarrow \infty} \mu_{i,j}$ is bounded. Therefore $\limsup_{N_t, K \rightarrow \infty} C_4$ is bounded as well. In order to prove convergence results, we examine the properties of the supremum over all the indices i, j .

$$\begin{aligned} \mathbb{E} \left[\sup_{i,j} \left| \Lambda_{i,j}(\mu_{i,j}) - \gamma_{i,j} \right|^k \right] &\stackrel{(a)}{\leq} \sum_{i,j} \mathbb{E} \left[\left| \Lambda_{i,j}(\mu_{i,j}) - \gamma_{i,j} \right|^k \right] \\ &\stackrel{(b)}{\leq} NK \frac{C_4}{N_t^{\frac{k}{2}}} = \frac{C_5}{N_t^{\frac{k}{2}-1}} \end{aligned} \quad (116)$$

where $C_5 = C_4N\beta$ (where $\beta = \frac{K}{N_t}$, a finite value). Inequality (a) follows from the linearity of expectation operation and (b) follows from the bound in (115). Finally, we make use of the following inequality (which can be shown easily, details omitted here)

$$\sup_{i,j} \mathbb{E} \left[\left| \Lambda_{i,j}(\mu_{i,j}) - \gamma_{i,j} \right|^k \right] \leq \mathbb{E} \left[\sup_{i,j} \left| \Lambda_{i,j}(\mu_{i,j}) - \gamma_{i,j} \right|^k \right]. \quad (117)$$

From (116) and (117) we deduce,

$$\sup_{i,j} \mathbb{E} \left[\left| \Lambda_{i,j}(\mu_{i,j}) - \gamma_{i,j} \right|^k \right] \leq \frac{C_5}{N_t^{\frac{k}{2}-1}}. \quad (118)$$

By taking k to be sufficiently high ($k \geq 6$ in this case), the right hand side of (118) is summable. By Markov Inequality ((5.31) of [32]) and the Borel Cantelli lemma (Theorem 4.3 of [32]), it follows that $\Lambda_{i,j}(\mu_{i,j}) - \gamma_{i,j} \xrightarrow[N_t, K \rightarrow \infty]{a.s.} 0$, $\forall i, j$.

APPENDIX E: CONVERGENCE PROOF OF THE DOWNLINK INTERFERENCE

We only focus on the convergence of inter-cell interference term (the convergence of the intra-cell interference follows in a similar manner). Throughout this section, we will denote $\Phi_n = \sum_{b=1}^N \sum_{k=1}^K \frac{\mu_{b,k}}{N_t} \mathbf{h}_{n,b,k} \mathbf{h}_{n,b,k}^H + \mathbf{I}$. Consider the inter-cell interference term $\sum_k |\mathbf{v}_{n,k}^H \mathbf{h}_{n,i,j}|^2$ ($n \neq i$). Substituting for $\mathbf{v}_{n,k} = \frac{\sqrt{\bar{\delta}_{n,k}}}{N_t} \Phi_n^{-1} \mathbf{h}_{n,n,k}$, we obtain,

$$\sum_k |\mathbf{v}_{n,k}^H \mathbf{h}_{n,i,j}|^2 = \sum_k \frac{\bar{\delta}_{n,k}}{N_t^2} \mathbf{h}_{n,i,j}^H \Phi_n^{-1} \mathbf{h}_{n,n,k} \mathbf{h}_{n,n,k}^H \Phi_n^{-1} \mathbf{h}_{n,i,j}. \quad (119)$$

Performing a change of variable $\mathbf{x}_{n,i,j} = \frac{1}{\sqrt{N_t}} \mathbf{h}_{n,i,j}$ and $\mathbf{x}_{n,n,k} = \frac{1}{\sqrt{N_t}} \mathbf{h}_{n,n,k}$, we obtain,

$$\begin{aligned} \sum_k |\mathbf{v}_{n,k}^H \mathbf{h}_{n,i,j}|^2 \\ = \mathbf{x}_{n,i,j}^H \Phi_n^{-1} \sum_k (\bar{\delta}_{n,k} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H) \Phi_n^{-1} \mathbf{x}_{n,i,j}. \end{aligned} \quad (120)$$

We use Lemma 8 to remove the column $\mathbf{x}_{n,i,j}$ from the matrix Φ_n . Denoting $\Phi'_n = \Phi_n - \mu_{i,j} \mathbf{x}_{n,i,j} \mathbf{x}_{n,i,j}^H$, we obtain

$$\begin{aligned} \sum_k |\mathbf{v}_{n,k}^H \mathbf{h}_{n,i,j}|^2 \\ = \frac{\mathbf{x}_{n,i,j}^H \Phi_n^{-1} \left(\sum_k \bar{\delta}_{n,k} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \right) \Phi_n^{-1} \mathbf{x}_{n,i,j}}{(1 + \mu_{i,j} \mathbf{x}_{n,i,j}^H \Phi_n^{-1} \mathbf{x}_{n,i,j})^2}. \end{aligned} \quad (121)$$

First, it can be easily verified that the denominator term in (121) converges to

$$(1 + \mu_{i,j} \mathbf{x}_{n,i,j}^H \Phi_n^{-1} \mathbf{x}_{n,i,j})^2 \asymp (1 + \mu_{i,j} \sigma_{n,i,j} \bar{m}_n)^2. \quad (122)$$

We now focus on the terms of the numerator of (121). Let us denote the matrix $\mathbf{A} = \Phi_n^{-1} \left(\sum_k \bar{\delta}_{n,k} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \right) \Phi_n^{-1}$. Applying Lemma 9 on the term $\mathbf{x}_{n,i,j}^H \mathbf{A} \mathbf{x}_{n,i,j}$, we have

$$\left| \mathbf{x}_{n,i,j}^H \mathbf{A} \mathbf{x}_{n,i,j} - \frac{\sigma_{n,i,j}}{N_t} \text{tr}(\mathbf{A}) \right|^m \leq \frac{C_m}{N_t^{\frac{m}{2}}} \|\mathbf{A}\|^m. \quad (123)$$

If $\|\mathbf{A}\| < \infty$, then we can conclude that

$$\mathbf{x}_{n,i,j}^H \mathbf{A} \mathbf{x}_{n,i,j} \asymp \frac{\sigma_{n,i,j}}{N_t} \text{tr}(\mathbf{A}). \quad (124)$$

The proof of the fact that $\|\mathbf{A}\|$ is bounded is deferred till the end of this section (see Subsection D at the end of this appendix).

Let us now focus on the term $\frac{1}{N_t} \text{tr}(\mathbf{A})$. First, we perform some straightforward manipulations.

$$\begin{aligned} \frac{1}{N_t} \text{tr}(\mathbf{A}) &= \frac{1}{N_t} \text{tr} \left(\Phi_n'^{-1} \left(\sum_k \bar{\delta}_{n,k} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \right) \Phi_n'^{-1} \right) \\ &\stackrel{(a)}{=} \frac{1}{N_t} \text{tr} \left(\left(\sum_k \bar{\delta}_{n,k} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \right) \Phi_n''^{-2} \right) \\ &\stackrel{(b)}{=} \frac{1}{N_t} \sum_k \bar{\delta}_{n,k} \left(\mathbf{x}_{n,n,k}^H \Phi_n''^{-2} \mathbf{x}_{n,n,k} \right), \end{aligned} \quad (125)$$

were (a) follows from $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ and (b) follows by noting $\text{tr}(\sum_i \mathbf{A}_i) = \sum_i \text{tr}(\mathbf{A}_i)$ and $\text{tr}(\mathbf{x}\mathbf{x}^H \mathbf{A}) = \mathbf{x}^H \mathbf{A} \mathbf{x}$. Applying Lemma 8 to extract the column vector $\mathbf{x}_{n,n,k}$ from the matrix $\Phi_n'^{-1}$, and denoting $\Phi_n'' = \Phi_n' - \mu_{n,k} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H$, we obtain,

$$\mathbf{x}_{n,n,k}^H \Phi_n'^{-2} \mathbf{x}_{n,n,k} = \frac{\mathbf{x}_{n,n,k}^H \Phi_n''^{-2} \mathbf{x}_{n,n,k}}{(1 + \mu_{n,k} \mathbf{x}_{n,n,k}^H \Phi_n'^{-1} \mathbf{x}_{n,n,k})^2}. \quad (126)$$

Note that the denominator term of (126) converges to the following:

$$(1 + \mu_{n,k} \mathbf{x}_{n,n,k}^H \Phi_n'^{-1} \mathbf{x}_{n,n,k})^2 \asymp (1 + \sigma_{n,n,k} \mu_{n,k} \bar{m}_n)^2. \quad (127)$$

Let us focus on the numerator term $\frac{1}{N_t} \sum_k \bar{\delta}_{n,k} \left(\mathbf{x}_{n,n,k}^H \Phi_n''^{-2} \mathbf{x}_{n,n,k} \right)$. Using the fact that $\frac{K}{N_t}$ is bounded, and applying the result of Theorem 5, eq. (69), it can be shown that (the exact details of the derivation are omitted)

$$\frac{\sum_k \bar{\delta}_{n,k} \left(\mathbf{x}_{n,n,k}^H \Phi_n''^{-2} \mathbf{x}_{n,n,k} \right)}{N_t} \asymp \frac{\sum_k \bar{\delta}_{n,k} \sigma_{n,n,k} \bar{m}'_n}{N_t}. \quad (128)$$

From the results of (121) and (128), it can be concluded that

$$\sum_k |\mathbf{v}_{n,k}^H \mathbf{h}_{n,i,j}|^2 \asymp \sum_k \frac{\bar{\delta}_{n,k} \bar{G}_{n,i,j} \bar{G}_{n,n,k} \bar{m}'_n}{N_t}, \quad n \neq i. \quad (129)$$

D. Boundedness of $\|\mathbf{A}\|$

We will complete the proof by showing that $\|\mathbf{A}\| < \infty$. First, it can be easily noticed that the matrix \mathbf{A} is Hermitian. We now show that the matrix \mathbf{A} is also positive semi-definite. For any vector $\mathbf{g} \in \mathbb{C}^{N_t \times 1}$, we examine $\mathbf{g}^H \mathbf{A} \mathbf{g}$.

$$\begin{aligned} \mathbf{g}^H \Phi_n'^{-1} \left(\sum_k \bar{\delta}_{n,k} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \right) \Phi_n'^{-1} \mathbf{g} \\ = \sum_k \bar{\delta}_{n,k} \mathbf{g}^H \Phi_n'^{-1} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \Phi_n'^{-1} \mathbf{g} \end{aligned} \quad (130)$$

Denoting $\mathbf{y} = \Phi_n'^{-1} \mathbf{g}$,

$$\sum_k \bar{\delta}_{n,k} \mathbf{g}^H \mathbf{y} \mathbf{y}^H \mathbf{g} = \sum_k \bar{\delta}_{n,k} |\mathbf{g}^H \mathbf{y}|^2 \geq 0 \quad (131)$$

since $\bar{\delta}_{n,k} \geq 0$ and $|\mathbf{g}^H \mathbf{y}|^2 \geq 0$. Therefore, the matrix \mathbf{A} is semi-definite. Next, we note that (since \mathbf{A} is Hermitian and positive semi-definite)

$$\|\mathbf{A}\| = \lambda_{\max}(\mathbf{A}). \quad (132)$$

Consider \mathbf{z} , the eigen vector corresponding to the maximum eigen value of the matrix \mathbf{A} .

$$\begin{aligned} \|\mathbf{A}\| &\leq \mathbf{z}^H \Phi_n'^{-1} \left(\sum_k \bar{\delta}_{n,k} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \right) \Phi_n'^{-1} \mathbf{z} \\ &= \sum_k \bar{\delta}_{n,k} \mathbf{z}^H \Phi_n'^{-1} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \Phi_n'^{-1} \mathbf{z} \\ &= \sum_k \left(\frac{\bar{\delta}_{n,k}}{\mu_{n,k}} \right) \mu_{n,k} \mathbf{z}^H \Phi_n'^{-1} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \Phi_n'^{-1} \mathbf{z} \end{aligned}$$

Note that $\beta = \max_{n,k} \frac{\bar{\delta}_{n,k}}{\mu_{n,k}}$ is a bounded value. Substituting, we have

$$\begin{aligned} \|\mathbf{A}\| &\leq \beta \sum_k \mu_{n,k} \mathbf{z}^H \Phi_n'^{-1} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \Phi_n'^{-1} \mathbf{z} \\ &\leq \beta \sum_k \mu_{n,k} \mathbf{z}^H \Phi_n'^{-1} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H \Phi_n'^{-1} \mathbf{z} + \beta \mathbf{z}^H \Phi_n''^{-2} \mathbf{z}. \end{aligned}$$

where the last inequality follows by noting that $\beta \mathbf{z}^H \Phi_n''^{-2} \mathbf{z} \geq 0$. Therefore,

$$\begin{aligned} \|\mathbf{A}\| &\leq \beta \mathbf{z}^H \Phi_n'^{-1} \left(\sum_k \mu_{n,k} \mathbf{x}_{n,n,k} \mathbf{x}_{n,n,k}^H + \mathbf{I} \right) \Phi_n'^{-1} \mathbf{z} \\ &= \beta \mathbf{z}^H \Phi_n'^{-1} \Phi_n' \Phi_n'^{-1} \mathbf{z} \\ &\leq \beta \mathbf{z}^H \Phi_n'^{-1} \mathbf{z} \leq \beta \lambda_{\max}(\Phi_n'^{-1}) < \infty. \end{aligned} \quad (133)$$

APPENDIX F: CONVERGENCE OF THE LAGRANGIAN

First we start with the proof of (40). Consider the Lagrangian given in the form of (3). For simplicity of notations, let us define

$$\begin{aligned} U_{i,j} &\triangleq \frac{|\mathbf{v}_{i,j}^H h_{i,i,j}|^2}{\gamma_{i,j}} \\ I_{i,j}^{(1)} &\triangleq \sum_{k \neq j} |\mathbf{v}_{i,k}^H \mathbf{h}_{i,i,j}|^2, \quad \bar{I}_{i,j}^{(1)} \triangleq \sum_{k \neq i} \frac{\bar{\delta}_{i,k} \bar{G}_{i,i,j} \bar{G}_{i,i,k} \bar{m}'_i}{N_t} \\ I_{i,j}^{(2)} &\triangleq \sum_{n,k} |\mathbf{v}_{n,k}^H \mathbf{h}_{n,i,j}|^2, \quad \bar{I}_{i,j}^{(2)} \triangleq \sum_{n,k} \frac{\bar{\delta}_{n,k} \bar{G}_{n,i,j} \bar{G}_{n,n,k} \bar{m}'_n}{N_t} \\ & \quad n \neq i. \end{aligned} \quad (134)$$

We examine the asymptotic convergence of the terms of the Lagrangian, i.e., $\lim_{N_t, K \rightarrow \infty} L(\mathbf{v}, \mu)$. By using arguments similar to the derivation of (128), the following can be shown:

$$\begin{aligned} \sum_{i,j} \frac{\mu_{i,j}}{N_t} \left[\frac{U_{i,j}}{\gamma_{i,j}} - I_{i,j}^{(1)} - \bar{I}_{i,j}^{(2)} - N_0 \right] \asymp \\ \sum_{i,j} \frac{\mu_{i,j}}{N_t} \left[\frac{\bar{U}_{i,j}}{\gamma_{i,j}} - \bar{I}_{i,j}^{(1)} - \bar{I}_{i,j}^{(2)} - N_0 \right]. \end{aligned} \quad (135)$$

Therefore,

$$\begin{aligned} \lim_{N_t, K \rightarrow \infty} L^{\text{DL}}(\mathbf{v}, \mu) \asymp \lim_{N_t, K \rightarrow \infty} \sum_{i,j} \mathbf{v}_{i,j}^H \mathbf{v}_{i,j} \\ + \sum_{i,j} \frac{\mu_{i,j}}{N_t} \left[\frac{\bar{U}_{i,j}}{\gamma_{i,j}} - \bar{I}_{i,j}^{(1)} - \bar{I}_{i,j}^{(2)} - N_0 \right]. \end{aligned} \quad (136)$$

Following the argument of (14), we note that

$$\frac{\bar{U}_{i,j}}{\gamma_{i,j}} - \bar{I}_{i,j}^{(1)} - \bar{I}_{i,j}^{(2)} - N_0 = 0 \quad \forall i, j. \quad (137)$$

From (135) and (137), we conclude that

$$\lim_{N_t, K \rightarrow \infty} L^{\text{DL}}(\mathbf{v}, \mu) \asymp \lim_{N_t, K \rightarrow \infty} \sum_{i,j} \mathbf{v}_{i,j}^H \mathbf{v}_{i,j}. \quad (138)$$

This completes the proof of (40).

Next, we proceed to the proof of (41). Consider the Lagrangian in the form of (4). Using (5), we can rewrite the term $\sum_{i,j} \mathbf{v}_{i,j}^H \mathbf{B}_{i,j} \mathbf{v}_{i,j}$ as

$$\begin{aligned} \sum_{i,j} \mathbf{v}_{i,j}^H \mathbf{B}_{i,j} \mathbf{v}_{i,j} &= \sum_{i,j} \left[\mathbf{v}_{i,j}^H \mathbf{v}_{i,j} - \frac{\mu_{i,j}}{\gamma_{i,j} N_t} |\mathbf{v}_{i,j}^H \mathbf{h}_{i,i,j}|^2 \right. \\ &\quad \left. + \sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k}}{N_t} |\mathbf{v}_{i,j}^H \mathbf{h}_{i,n,k}|^2 \right]. \end{aligned} \quad (139)$$

Substituting for $\mathbf{v}_{i,j} = \sqrt{\frac{\bar{\delta}_{i,j}}{N_t}} \hat{\mathbf{v}}_{i,j}$ in (139), we obtain

$$\begin{aligned} \sum_{i,j} \mathbf{v}_{i,j}^H \mathbf{B}_{i,j} \mathbf{v}_{i,j} &= \sum_{i,j} \frac{\bar{\delta}_{i,j}}{N_t} \left[\hat{\mathbf{v}}_{i,j}^H \hat{\mathbf{v}}_{i,j} \right. \\ &\quad \left. - \frac{\mu_{i,j}}{\gamma_{i,j} N_t} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,i,j}|^2 + \sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k}}{N_t} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,n,k}|^2 \right]. \end{aligned} \quad (140)$$

We now examine the asymptotic convergence of the terms on the right hand side of (140). The following convergence results can be shown (the details are omitted here as they follow from steps similar to the proof of previous results).

$$\hat{\mathbf{v}}_{i,j}^H \hat{\mathbf{v}}_{i,j} \asymp \frac{\sigma_{i,i,j} \bar{m}'_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2}. \quad (141)$$

$$\frac{\mu_{i,j}}{N_t} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,i,j}|^2 \asymp \mu_{i,j} \left(\frac{\sigma_{i,i,j} \bar{m}_i}{1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i} \right)^2. \quad (142)$$

$$\begin{aligned} \sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k}}{N_t} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,n,k}|^2 &\asymp \\ \sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k} \sigma_{i,i,j} \sigma_{i,n,k} \bar{m}'_i}{N_t (1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2 (1 + \mu_{n,k} \sigma_{i,n,k} \bar{m}_i)^2}. \end{aligned} \quad (143)$$

From (142) and using the fact that $\gamma_{i,i,j} = \sigma_{i,i,j} \mu_{i,j} \bar{m}_i$, it follows that

$$\frac{\mu_{i,j}}{\gamma_{i,j} N_t^2} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,i,j}|^2 \asymp \frac{\sigma_{i,i,j} \bar{m}_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2}. \quad (144)$$

Additionally, similar to the proof of (135), it can be shown that

$$\begin{aligned} \sum_{i,j} \frac{\bar{\delta}_{i,j}}{N_t} \left[\frac{1}{N_t} \hat{\mathbf{v}}_{i,j}^H \hat{\mathbf{v}}_{i,j} - \frac{\mu_{i,j}}{\gamma_{i,j} N_t^2} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,i,j}|^2 \right. \\ \left. + \sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k}}{N_t^2} |\hat{\mathbf{v}}_{i,j}^H \mathbf{h}_{i,n,k}|^2 \right] \asymp \\ \sum_{i,j} \frac{\bar{\delta}_{i,j}}{N_t} \left[\frac{\sigma_{i,i,j} \bar{m}'_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2} - \frac{\sigma_{i,i,j} \bar{m}_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2} \right. \\ \left. + \sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k} \sigma_{i,i,j} \sigma_{i,n,k} \bar{m}'_i}{N_t (1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2 (1 + \mu_{n,k} \sigma_{i,n,k} \bar{m}_i)^2} \right]. \end{aligned} \quad (145)$$

Therefore, we conclude that

$$\begin{aligned} \lim_{N_t, K \rightarrow \infty} L^{\text{UL}}(\mathbf{v}, \mu) &\asymp \lim_{N_t, K \rightarrow \infty} \sum_{i,j} \frac{\mu_{i,j} N_0}{N_t} \\ &\quad + \sum_{i,j} \frac{\bar{\delta}_{i,j}}{N_t} \left[\frac{\sigma_{i,i,j} \bar{m}'_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2} - \frac{\sigma_{i,i,j} \bar{m}_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2} \right. \\ &\quad \left. + \sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k} \sigma_{i,i,j} \sigma_{i,n,k} \bar{m}'_i}{N_t (1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2 (1 + \mu_{n,k} \sigma_{i,n,k} \bar{m}_i)^2} \right]. \end{aligned} \quad (146)$$

Finally, we show the following result:

$$\begin{aligned} \frac{\sigma_{i,i,j} \bar{m}'_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2} - \frac{\sigma_{i,i,j} \bar{m}_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2} \\ + \sum_{(n,k) \neq (i,j)} \frac{\mu_{n,k} \sigma_{i,i,j} \sigma_{i,n,k} \bar{m}'_i}{N_t (1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2 (1 + \mu_{n,k} \sigma_{i,n,k} \bar{m}_i)^2} = 0. \end{aligned} \quad (147)$$

The proof of this result is provided next. Consider the sum of the first and last terms of (147).

$$\begin{aligned} \sum_{n,k} \frac{\mu_{n,k} \sigma_{i,i,j} \sigma_{i,n,k} \bar{m}'_i}{N_t (1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2 (1 + \mu_{n,k} \sigma_{i,n,k} \bar{m}_i)^2} \\ + \frac{\sigma_{i,i,j} \bar{m}'_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2} \\ = \frac{\sigma_{i,i,j} \bar{m}'_i}{(1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2} \left(1 + \frac{1}{N_t} \sum_{n,k} \frac{\mu_{n,k} \sigma_{i,n,k}}{(1 + \mu_{n,k} \sigma_{i,n,k} \bar{m}_i)^2} \right). \end{aligned} \quad (148)$$

Multiplying and dividing the right hand side of (148) by \bar{m}_i , we obtain

$$(148) = \frac{\sigma_{i,i,j} \bar{m}'_i}{\bar{m}_i (1 + \mu_{i,j} \sigma_{i,i,j} \bar{m}_i)^2} \left(\bar{m}_i + \frac{1}{N_t} \sum_{n,k} \frac{\mu_{n,k} \sigma_{i,n,k} \bar{m}_i}{(1 + \mu_{n,k} \sigma_{i,n,k} \bar{m}_i)^2} \right). \quad (149)$$

Recall the fixed point equation for the computation of \bar{m}_i in (12). Upon rearranging the terms

$$\bar{m}_i = 1 - \frac{1}{N_t} \sum_{n,k} \frac{\sigma_{i,n,k} \mu_{n,k} \bar{m}_i}{1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i}. \quad (150)$$

Using \bar{m}_i for the terms inside the brackets of the right hand side of (149), we have,

$$\begin{aligned} 1 - \frac{1}{N_t} \sum_{n,k} \frac{\sigma_{i,n,k} \mu_{n,k} \bar{m}_i}{1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i} + \frac{1}{N_t} \sum_{n,k} \frac{\mu_{n,k} \sigma_{i,n,k} \bar{m}_i}{(1 + \mu_{n,k} \sigma_{i,n,k} \bar{m}_i)^2} \\ = 1 - \frac{1}{N_t} \sum_{n,k} \frac{(\sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}. \end{aligned} \quad (151)$$

Rearranging the expression for \bar{m}'_i in (70), we have,

$$\frac{\bar{m}'_i}{\bar{m}_i} = 1 - \frac{1}{N_t} \sum_{n,k} \left(\frac{\sigma_{i,n,k} \mu_{n,k} \bar{m}_i}{1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i} \right)^2. \quad (152)$$

Using (151) and (152) in (149), it can be verified that

$$(149) = \frac{\sigma_{i,i,j} \bar{m}_i}{(1 + \sigma_{i,n,k} \mu_{n,k} \bar{m}_i)^2}. \quad (153)$$

From (153), the result of (147) follows.

Finally, using (147) in (146) we conclude that

$$\lim_{N_t, K \rightarrow \infty} L^{\text{UL}}(\mathbf{v}, \boldsymbol{\mu}) \asymp \lim_{N_t, K \rightarrow \infty} \sum_{i,j} \frac{\mu_{i,j} N_0}{N_t}. \quad (154)$$

This completes the proof of (41).

APPENDIX G: CONVERGENCE OF UPLINK POWER ALLOCATION

The proof proceeds by an approach similar to [11] of finding the upper and lower bounds of the optimal solution and show that these bounds coincides asymptotically with the solution obtained by the ROBF algorithm. Recall the fixed point equation for the computation of the uplink power allocation in the CBF algorithm in (7). It can be rewritten as

$$f_{i,j}(\boldsymbol{\beta}) = \frac{1}{\frac{1}{\gamma_{i,j} N_t} \mathbf{h}_{i,i,j}^H (\boldsymbol{\Sigma}_i^{\boldsymbol{\beta}} + \mathbf{I}_{N_t})^{-1} \mathbf{h}_{i,i,j}} \quad \forall i, j \quad (155)$$

Also recall from the arguments of [9] that $f_{i,j}(\boldsymbol{\beta})$ is a standard function.

Let us denote $\mu_{i,j}(\delta_{i,j})$, as the solution provided by the ROBF algorithm with target SINR $\gamma_{i,j} + \delta_{i,j} \quad \forall i, j$, where $\delta_{i,j} \geq 0$ is a small positive constant (i.e. the solution provided by the fixed point equation (11) with $\gamma_{i,j}$ replaced by $\gamma_{i,j} + \delta_{i,j}$, $\forall i, j$). We now examine the achieved SINR in the uplink with a power allocation of $\mu_{i,j}(\delta_{i,j})$, i.e.,

$$\begin{aligned} \Lambda_{i,j}(\boldsymbol{\mu}(\boldsymbol{\delta})) \\ = \frac{\mu_{i,j}(\delta_{i,j})}{N_t} \mathbf{h}_{i,i,j}^H (\boldsymbol{\Sigma}_i^{\mu_{i,j}(\delta_{i,j})} + \alpha_i \mathbf{I}_{N_t})^{-1} \mathbf{h}_{i,i,j}. \end{aligned} \quad (156)$$

Similar to the result of Theorem 3 (convergence of the uplink SINR), it can be proved that

$$\Lambda_{i,j}(\mu_{i,j}(\delta_{i,j})) - (\gamma_{i,j} + \delta_{i,j}) \xrightarrow[N_t, K \rightarrow \infty]{a.s.} 0, \quad \forall i, j. \quad (157)$$

Since $\delta_{i,j} \geq 0$, we have

$$\Lambda_{i,j}(\mu_{i,j}(\delta_{i,j})) - \gamma_{i,j} \stackrel{a.s.}{\geq} 0, \quad \forall i, j \quad (158)$$

Since the achieved SINR with an uplink power $\mu_{i,j}(\delta_{i,j})$ allocation is asymptotically greater than $\gamma_{i,j}$, the power allocation $\mu_{i,j}(\delta_{i,j})$ is a feasible solution to the uplink problem with target SINR $\gamma_{i,j}$. From the monotonically property of the standard function $f(\boldsymbol{\beta})$, we can conclude that

$$\mu_{i,j}(\delta_{i,j}) \geq \lambda_{i,j} \quad \forall i, j. \quad (159)$$

Similarly, let us define $\mu_{i,j}(-\delta_{i,j})$ as the solution to the fixed point equation (11) with target SINR $\gamma_{i,j} - \delta_{i,j}$, $\forall i, j$. By similar argument as above, we have

$$\Lambda_{i,j}(\mu_{i,j}(-\delta_{i,j})) - (\gamma_{i,j} - \delta_{i,j}) \xrightarrow[N_t, K \rightarrow \infty]{a.s.} 0, \quad \forall i, j. \quad (160)$$

Using the definition of almost sure convergence, we can write that for $\epsilon_{i,j} > 0$, there exists N' such that for all $N_t > N'$, we can have

$$|\Lambda_{i,j}(\mu_{i,j}(-\delta_{i,j})) - (\gamma_{i,j} - \delta_{i,j})| < \epsilon_{i,j}. \quad (161)$$

If we can select a value of $\epsilon_{i,j} < \delta_{i,j}$, then there exists a large enough N'' such that for all $N_t > N''$, the following

$$\frac{\mu_{i,j}(-\delta_{i,j})}{N_t} \mathbf{h}_{i,i,j}^H (\boldsymbol{\Sigma}_i^{\mu_{i,j}(-\delta_{i,j})} + \mathbf{I}_{N_t})^{-1} \mathbf{h}_{i,i,j} \leq \gamma_{i,j}, \quad \forall i, j \quad (162)$$

holds true. Equation (162) implies that

$$\begin{aligned} \mu_{i,j}(-\delta_{i,j}) &\leq \frac{1}{\frac{1}{\gamma_{i,j} N_t} \mathbf{h}_{i,i,j}^H (\boldsymbol{\Sigma}_i^{\mu_{i,j}(-\delta_{i,j})} + \mathbf{I}_{N_t})^{-1} \mathbf{h}_{i,i,j}}, \quad \forall i, j. \\ &= f_{i,j}(\boldsymbol{\mu}(-\boldsymbol{\delta})). \end{aligned} \quad (163)$$

In other words, $\boldsymbol{\mu}(-\boldsymbol{\delta})$ an infeasible point for the standard function $f_{i,j}(\boldsymbol{\mu}(-\boldsymbol{\delta}))$.

Let us consider that we want to find a fixed point to

$$\beta_{i,j} = f_{i,j}(\boldsymbol{\beta}) \quad \forall i, j. \quad (164)$$

The fixed point corresponding to this is the optimal uplink power allocation of the CBF algorithm $\lambda_{i,j}$. We start with an initial value of $\beta_{i,j}^0 = \mu_{i,j}(-\delta_{i,j})$. Let us consider the iterations of $\beta_{i,j}^{t+1} = f_{i,j}(\boldsymbol{\beta}^t)$, with $\boldsymbol{\beta}^0 = \boldsymbol{\mu}(-\boldsymbol{\delta})$. Rewriting the first iteration, we have,

$$\beta_{i,j}^1 = f(\boldsymbol{\mu}(-\boldsymbol{\delta})) \stackrel{(a)}{\geq} \mu_{i,j}(-\delta_{i,j}) \quad \forall i, j, \quad (165)$$

where (a) follows from (163). Also note that this is true for all i, j . Using the fact that the function f is standard and hence monotonic, the sequence $\beta_{i,j}^t$, $t = 0, 1, 2, \dots$, monotonically increases and converges to $\lambda_{i,j}$. Therefore, it follows that

$$\mu_{i,j}(-\delta_{i,j}) \leq \lambda_{i,j}. \quad (166)$$

Therefore, from the above arguments, we conclude that

$$\mu_{i,j}(-\delta_{i,j}) \stackrel{a.s.}{\leq} \lambda_{i,j} \stackrel{a.s.}{\leq} \mu_{i,j}(\delta_{i,j}) \quad \forall i, j. \quad (167)$$

By taking $\delta_{i,j}$ arbitrarily small, we have

$$\mu_{i,j} - \lambda_{i,j} \xrightarrow[N_t, K \rightarrow \infty]{a.s.} 0 \quad \forall i, j. \quad (168)$$

Similar to the proofs as before,

$$\frac{\sum_{i,j} \mu_{i,j}}{N_t} - \frac{\sum_{i,j} \lambda_{i,j}}{N_t} \xrightarrow[N_t, K \rightarrow \infty]{a.s.} 0 \quad \forall i, j. \quad (169)$$

APPENDIX H: IMPACT OF IMPERFECT CSI AND PILOT CONTAMINATION

Recall the uplink receive filter with CSI estimate in (50). First note that for the MMSE estimate of the form (47), $\hat{\mathbf{h}}_{i,n,k}$, $n \neq i$ and $\hat{\mathbf{h}}_{i,i,k}$ are related as

$$\hat{\mathbf{h}}_{i,n,k} = \frac{\sigma_{i,n,k}}{\sigma_{i,i,k}} \hat{\mathbf{h}}_{i,i,k} \quad (170)$$

Using (170) in (51), we have

$$\begin{aligned} \Psi_i &= \sum_{n=1}^N \sum_{k=1}^K \mu_{n,k} \hat{\mathbf{h}}_{i,n,k} \hat{\mathbf{h}}_{i,n,k}^H + \mathbf{I} \\ &= \sum_{k=1}^K \left(\frac{\sum_{n=1}^N \sigma_{i,n,k} \mu_{n,k}}{\sigma_{i,i,k} N_t} \right) \hat{\mathbf{h}}_{i,i,k} \hat{\mathbf{h}}_{i,i,k}^H + \mathbf{I} \\ &= \sum_{k=1}^K \xi_{i,k} \hat{\mathbf{x}}_{i,i,k} \hat{\mathbf{x}}_{i,i,k}^H + \mathbf{I} \end{aligned} \quad (171)$$

where

$$\xi_{i,k} = \frac{\sum_{n=1}^N \sigma_{i,n,k} \mu_{n,k}}{\sigma_{i,i,k}}. \quad (172)$$

Additionally, let us denote

$$\begin{aligned} \Psi_{i,j} &= \sum_{k=1, k \neq j}^K \xi_{i,k} \hat{\mathbf{x}}_{i,i,k} \hat{\mathbf{x}}_{i,i,k}^H + \mathbf{I}_{N_t} \\ \Psi_{i,j,l} &= \sum_{k=1, k \neq j, k \neq l}^K \xi_{i,k} \hat{\mathbf{x}}_{i,i,k} \hat{\mathbf{x}}_{i,i,k}^H + \mathbf{I}_{N_t}. \end{aligned} \quad (173)$$

First, we focus on the asymptotic equivalent of the useful signal term.

$$\begin{aligned} \mathbf{v}_{i,j}^{\text{est}H} \mathbf{h}_{i,i,j} &= \frac{\sqrt{\bar{\delta}_{i,j}}}{N_t} \hat{\mathbf{h}}_{i,i,j}^H \Psi_i^{-1} \mathbf{h}_{i,i,j} \\ &= \sqrt{\bar{\delta}_{i,j}} \hat{\mathbf{x}}_{i,i,j}^H \Psi_i^{-1} \mathbf{x}_{i,i,j} \end{aligned} \quad (174)$$

First note that using Lemma 8, we can remove the column $\hat{\mathbf{x}}_{i,i,j}$ from the matrix Ψ_i^{-1} as follows:

$$\begin{aligned} (174) &= \frac{\sqrt{\bar{\delta}_{i,j}} \hat{\mathbf{x}}_{i,i,j}^H \Psi_i^{-1} \mathbf{x}_{i,i,j}}{1 + \xi_{i,j} \hat{\mathbf{x}}_{i,i,j}^H \Psi_{i,j}^{-1} \hat{\mathbf{x}}_{i,i,j}} \\ &\asymp \frac{\sqrt{\bar{\delta}_{i,j}} \hat{\sigma}_{i,i,j} \bar{m}_i^{\text{est}}}{1 + \xi_{i,j} \hat{\sigma}_{i,i,j} \bar{m}_i^{\text{est}}}, \end{aligned} \quad (175)$$

where \bar{m}_i^{est} can be computed as in (54). The result in (175) can be derived as follows: First, we look at the numerator term

$$\begin{aligned} &\hat{\mathbf{x}}_{i,i,j}^H \Psi_{i,j}^{-1} \mathbf{x}_{i,i,j} \\ &= \sigma'_{i,i,j} \left(\sum_{b=1}^N \mathbf{x}_{i,b,j} + \frac{\mathbf{n}_{\text{Tr}}}{\sqrt{P_{\text{Tr}}}} \right)^H \Psi_{i,j}^{-1} \mathbf{x}_{i,i,j} \\ &\stackrel{(a)}{\asymp} \sigma'_{i,i,j} \sigma_{i,i,j} \bar{m}_i^{\text{est}} \stackrel{(b)}{=} \hat{\sigma}_{i,i,j} \bar{m}_i^{\text{est}}, \end{aligned} \quad (176)$$

where (a) follows from the results of Lemma 9, Theorem 5, and noting that the column vectors $\mathbf{x}_{i,b,j} \forall b \neq i$, \mathbf{n}_{Tr} are independent of $\mathbf{x}_{i,i,j}$. Step (b) follows using the relation in (49). Similarly, the denominator term converges to

$$1 + \xi_{i,j} \hat{\mathbf{x}}_{i,i,j}^H \Psi_{i,j}^{-1} \hat{\mathbf{x}}_{i,i,j} \asymp 1 + \xi_{i,j} \hat{\sigma}_{i,i,j} \bar{m}_i^{\text{est}}.$$

We now investigate the asymptotic equivalent of the sum of interference power at $\text{UT}_{i,j}$ from the BS of cell n , given by

$$\begin{aligned} \sum_k |\mathbf{v}_{n,k}^{\text{est}H} \mathbf{h}_{n,i,j}|^2 &= \sum_k \frac{\bar{\delta}_{n,k}}{N_t^2} \mathbf{h}_{n,i,j}^H \Psi_n^{-1} \hat{\mathbf{h}}_{n,n,k} \hat{\mathbf{h}}_{n,n,k}^H \Psi_n^{-1} \mathbf{h}_{n,i,j} \\ &= \sum_k \bar{\delta}_{n,k} \mathbf{x}_{n,i,j} \Psi_n^{-1} \hat{\mathbf{x}}_{n,n,k} \hat{\mathbf{x}}_{n,n,k}^H \Psi_n^{-1} \mathbf{x}_{n,i,j}. \end{aligned}$$

We note that the CSI estimate of $\text{UT}_{i,j}$ is contaminated from the pilot signal of $\text{UT}_{n,j}$, ($n \neq i$). Therefore, we first analyze the interference signal from $\text{UT}_{n,j}$, ($n \neq i$), as follows:

$$\begin{aligned} &\bar{\delta}_{n,j} \mathbf{x}_{n,i,j}^H \Psi_n^{-1} \hat{\mathbf{x}}_{n,n,j} \hat{\mathbf{x}}_{n,n,j}^H \Psi_n^{-1} \mathbf{x}_{n,i,j} \\ &\stackrel{(a)}{=} \frac{\bar{\delta}_{n,j} \hat{\mathbf{x}}_{n,n,j}^H \Psi_n^{-1} \mathbf{x}_{n,i,j} \mathbf{x}_{n,i,j}^H \Psi_n^{-1} \hat{\mathbf{x}}_{n,n,j}}{(1 + \xi_{n,j} \hat{\mathbf{x}}_{n,n,j}^H \Psi_{n,j}^{-1} \hat{\mathbf{x}}_{n,n,j})^2} \\ &\stackrel{(b)}{\asymp} \frac{\bar{\delta}_{n,j} (\sigma_{n,i,j} \sigma'_{n,n,j} \bar{m}_n^{\text{est}})^2}{(1 + \xi_{n,j} \hat{\sigma}_{n,n,j} \bar{m}_n^{\text{est}})^2} \end{aligned} \quad (177)$$

where in (a), we have removed the column $\hat{\mathbf{x}}_{n,n,j}^H$ from the matrix Ψ_n^{-1} using Lemma 8, and the result of (b) is derived similar to (175).

Next, we consider the sum of interference signals from rest of the UTs in cell n , i.e.,

$$\begin{aligned} &\sum_{k \neq j} \bar{\delta}_{n,k} \mathbf{x}_{n,i,j}^H \Psi_n^{-1} \hat{\mathbf{x}}_{n,n,k} \hat{\mathbf{x}}_{n,n,k}^H \Psi_n^{-1} \mathbf{x}_{n,i,j} \\ &= \mathbf{x}_{n,i,j} \Psi_n^{-1} \mathbf{B}_{n,j} \Psi_n^{-1} \mathbf{x}_{n,i,j}, \end{aligned} \quad (178)$$

where $\mathbf{B}_{n,j} = \sum_{k \neq j} \bar{\delta}_{n,k} \hat{\mathbf{x}}_{n,n,k} \hat{\mathbf{x}}_{n,n,k}^H$. Using Lemma 8, we can decompose Ψ_n^{-1} as

$$\Psi_n^{-1} = \Psi_{n,j}^{-1} - \frac{\xi_{n,j} \Psi_{n,j}^{-1} \hat{\mathbf{x}}_{n,n,j} \hat{\mathbf{x}}_{n,n,j}^H \Psi_{n,j}^{-1}}{1 + \xi_{n,j} \hat{\mathbf{x}}_{n,n,j}^H \Psi_{n,j}^{-1} \hat{\mathbf{x}}_{n,n,j}}. \quad (179)$$

Using (179) in (178), we obtain

$$\begin{aligned} &\mathbf{x}_{n,i,j}^H \Psi_n^{-1} \mathbf{B}_{n,j} \Psi_n^{-1} \mathbf{x}_{n,i,j} \\ &= \mathbf{x}_{n,i,j}^H \Psi_{n,j}^{-1} \mathbf{B}_{n,j} \Psi_{n,j}^{-1} \mathbf{x}_{n,i,j} \end{aligned} \quad (180)$$

$$\begin{aligned} &+ \frac{\xi_{n,j}^2 (\mathbf{x}_{n,i,j}^H \Psi_{n,j}^{-1} \hat{\mathbf{x}}_{n,n,j})^2 \hat{\mathbf{x}}_{n,n,j}^H \Psi_{n,j}^{-1} \mathbf{B}_{n,j} \Psi_{n,j}^{-1} \hat{\mathbf{x}}_{n,n,j}}{(1 + \xi_{n,j} \hat{\mathbf{x}}_{n,n,j}^H \Psi_{n,j}^{-1} \hat{\mathbf{x}}_{n,n,j})^2} \\ &- 2\text{Re} \left\{ \frac{\xi_{n,j} (\hat{\mathbf{x}}_{n,n,j}^H \Psi_{n,j}^{-1} \mathbf{x}_{n,i,j}) \mathbf{x}_{n,i,j}^H \Psi_{n,j}^{-1} \mathbf{B}_{n,j} \Psi_{n,j}^{-1} \hat{\mathbf{x}}_{n,n,j}}{1 + \xi_{n,j} \hat{\mathbf{x}}_{n,n,j}^H \Psi_{n,j}^{-1} \hat{\mathbf{x}}_{n,n,j}} \right\}. \end{aligned} \quad (181)$$

Consider the term (180). It can be proved that the spectral norm of $\Psi_{n,j}^{-1} \mathbf{B}_{n,j} \Psi_{n,j}^{-1}$ is bounded (following the steps in (133)). Therefore, using

$$\mathbf{x}_{n,i,j} \Psi_{n,j}^{-1} \mathbf{B}_{n,j} \Psi_{n,j}^{-1} \mathbf{x}_{n,i,j} \asymp \frac{\sigma_{n,i,j}}{N_t} \text{tr}(\Psi_{n,j}^{-1} \mathbf{B}_{n,j} \Psi_{n,j}^{-1}). \quad (183)$$

Using some straightforward steps (similar to the derivation of (125)), it can be shown that

$$\frac{1}{N_t} \text{tr}(\Psi_{n,j}^{-1} \mathbf{B}_{n,j} \Psi_{n,j}^{-1}) = \frac{1}{N_t} \sum_{k \neq j} \bar{\delta}_{n,k} \hat{\mathbf{x}}_{n,n,k}^H \Psi_{n,j}^{-2} \hat{\mathbf{x}}_{n,n,k}. \quad (184)$$

The right hand side of (184) can be further analyzed as follows:

$$\begin{aligned} (184) &= \frac{1}{N_t} \sum_{k \neq j} \frac{\bar{\delta}_{n,k} \hat{\mathbf{x}}_{n,n,k}^H \Psi_{n,j}^{-2} \hat{\mathbf{x}}_{n,n,k}}{(1 + \xi_{n,k} \hat{\mathbf{x}}_{n,n,k}^H \Psi_{n,j,k}^{-1} \hat{\mathbf{x}}_{n,n,k})^2} \\ &\asymp \frac{1}{N_t} \sum_{k \neq j} \frac{\bar{\delta}_{n,k} \hat{\sigma}_{n,n,k} (\bar{m}'_n)^{\text{est}}}{(1 + \xi_{n,k} \hat{\sigma}_{n,n,k} \bar{m}_n^{\text{est}})^2}, \end{aligned} \quad (185)$$

where $(\bar{m}'_n)^{\text{est}}$ can be evaluated as in (55), and the result of (185) has been derived following the approach of (128). Similarly, it can be shown that

$$(181) \asymp \sum_{k \neq j} \frac{\bar{\delta}_{n,k} \xi_{n,j}^2 (\sigma_{n,i,j} \sigma'_{n,n,j} \bar{m}_n^{\text{est}})^2 \hat{\sigma}_{n,n,j} \hat{\sigma}_{n,n,k} (\bar{m}'_n)^{\text{est}}}{N_t (1 + \xi_{n,j} \hat{\sigma}_{n,n,j} \bar{m}_n^{\text{est}})^2 (1 + \xi_{n,k} \hat{\sigma}_{n,n,k} \bar{m}_n^{\text{est}})^2},$$

$$(182) \asymp -2 \sum_{k \neq j} \frac{\sigma_{n,i,j}^2 (\sigma'_{n,n,j})^2 \bar{m}_n^{\text{est}} \hat{\sigma}_{n,n,k} (\bar{m}'_n)^{\text{est}}}{N_t (1 + \xi_{n,j} \hat{\sigma}_{n,n,j} \bar{m}_n^{\text{est}}) (1 + \xi_{n,k} \hat{\sigma}_{n,n,k} \bar{m}_n^{\text{est}})^2}.$$

Combining the convergence results of the useful signal and the interference signal terms derived in this Appendix, we obtain (56).

REFERENCES

- [1] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [2] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [3] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [5] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [6] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Designing multi-user mimo for energy efficiency: When is massive mimo the answer?" in *Proc. IEEE Wireless Communications and Networking Conference*, Apr. 2014.
- [7] A. Liu and V. Lau, "Hierarchical interference mitigation for massive MIMO cellular networks," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4786–4797, Sep. 2014.
- [8] H. Huh, G. Caire, H. Papadopoulos, and S. Ramprasad, "Achieving "Massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, Sep. 2012.
- [9] 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.
- [10] S. Lakshminaryana, J. Hoydis, M. Debbah, and M. Assaad, "Asymptotic Analysis of Distributed Multi-cell Beamforming," in *International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC'10)*, Sep. 2010, pp. 2105–2110.
- [11] R. Zakhour and S. Hanly, "Base Station Cooperation on the Downlink: Large System Analysis," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2079–2106, Apr. 2012.
- [12] —, "Min-max power allocation in cellular networks with coordinated beamforming," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 287–302, Feb. 2013.
- [13] V. Nguyen and J. Evans, "Multiuser transmit beamforming via regularized channel inversion : A large system analysis," in *IEEE Global Telecommunications Conference, 2008*, Nov 2008, pp. 1–4.
- [14] S. Wagner, R. Couillet, M. Debbah, and D. Slock, "Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels Under Limited Feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.
- [15] A. Tulino and S. Verdú, "Asymptotic analysis of improved linear receivers for BPSK-CDMA subject to fading," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1544–1555, Aug. 2001.
- [16] H. Huh, A. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 291–2934, May 2012.
- [17] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO in real propagation environments," *IEEE Trans. Wireless Commun.*, 2014, submitted. [Online]. Available: <http://arxiv.org/abs/1403.3376>
- [18] A. Wiesel, Y. Eldar, and S. Shamai, "Linear Precoding via Conic Optimization for Fixed MIMO Receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
- [19] I. Viering, H. Hofstetter, and W. Utschick, "Spatial long-term variations in urban, rural and indoor environments," in *Proc. COST273 5th Meeting, Lisbon, Portugal, 2002*.
- [20] A. Kammoun, A. Müller, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: Reducing complexity in massive MIMO," *IEEE Trans. Signal Process.*, 2014, submitted. [Online]. Available: <http://arxiv.org/pdf/1310.1806.pdf>
- [21] D. M. Young, *Iterative Solutions of Large Linear Systems*. Academic Press, New York, 1971.
- [22] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Proc. IEEE International Symposium on Information Theory Proceedings (ISIT), 2012*, Jul. 2012, pp. 1137–1141.
- [23] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Upper Saddle River, NJ, USA, 1993.
- [24] J. E. Rombaldi, *Analyse Matricielle*. EDP Sciences, 2000.
- [25] J. W. Silverstein and Z. Bai, "On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.
- [26] Z. Bai and J. W. Silverstein, "No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large Dimensional Sample Covariance Matrices," *Annals of Probability*, vol. 26, pp. 316–345, Jan. 1998.
- [27] M. Peacock, I. Collings, and M. Honig, "Eigenvalue Distributions of Sums and Products of Large Random Matrices Via Incremental Matrix Expansions," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, May 2008.
- [28] R. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.
- [29] M. Chiang and P. Hande and T. Lan and C. W. Tan, *Power Control in Wireless Cellular Networks*. Foundations and Trends in Networking, NOW Publishers, 2008.
- [30] J. Zander and M. Frodigh, "Comment on "Performance of optimum transmitter power control in cellular radio systems"," *IEEE Trans. Veh. Technol.*, vol. 43, no. 3, p. 636, Aug. 1994.
- [31] W. Hachem, P. Loubaton, and J. Najim, "Deterministic Equivalents for Certain Functionals of Large Random Matrices," *Annals of Applied Probability*, vol. 17, no. 3, pp. 875–930, 2007.
- [32] P. Billingsley, *Probability and Measure*. John Wiley & Sons, Inc., Hoboken, NJ, 1995.