Boriana Boiadjieva and Anja Klein, ”Distributed Two-Stage Beamforming with Power Allocation in Multicell Massive MIMO,” in *Proc. of the IEEE International Conference on Communications - (IEEE ICC 2019)*, Shanghai, China, Mai 2019.

Distributed Two-Stage Beamforming with Power Allocation in Multicell Massive MIMO

Boriana Boiadjieva and Anja Klein
Communications Engineering Lab
Technische Universität Darmstadt, Merckstrasse 25, 64283 Darmstadt, Germany
{b.boiadjieva, a.klein}@nt.tu-darmstadt.de

***Abstract*—We consider a multicellular system in the downlink where at each base station (BS), massive multiple-input multiple-output (MIMO) is utilized via large antenna array. However, big numbers of antennas and users result in a huge channel state information (CSI) overhead and big computational complexity. To cope with these challenges, we propose a two-stage beamforming with power allocation using the signal to leakage and noise ratio (SLNR) as a performance metric since it is known to decouple the optimization problems compared to conventional design methods, allowing independent, distributed processing at each BS. As a first contribution, we derive a deterministic equivalent of the SLNR in the distributed two-stage beamforming context using random matrix theory which provides very accurate approximations in closed-form. The second contribution consists of beamforming design and power allocation. More precisely, in the first stage of the proposed beamforming, an outer beamformer is designed using the deterministic equivalent of the SLNR which requires only statistical CSI and produces an effective system of lower dimension. In the second stage, an inner beamformer applies regularized zero forcing on the low-dimensional effective channel to combat the interference. Additionally, to improve the user fairness, power allocation which maximizes the minimum effective SLNR is proposed and shown to be a convex problem. Simulation results confirm that both the proposed beamforming as well as the combination of the proposed beamforming with power allocation reduce the system dimensionality without degradation of system performance in terms of throughput and user fairness.**

I. Introduction

The multiple-input multiple-output (MIMO) technology is a fundamental part of current and future wireless systems. It has already been adopted in the standards for wireless communications since it offers data rate increase and improved link reliability without the need of additional investments in transmit power or bandwidth [1], [2]. Over the last decade, researchers have intensively studied multiuser MIMO where several users are served simultaneously in the same time-frequency resource using beamforming, to increase the spectral efficiency and the throughput [3]. More recently, the focus of the research community has moved to multiuser massive MIMO where a large number of antennas is used in order to meet the future capacity demand. However, the application of large antenna arrays especially in multicellular context tremendously increases the channel state information (CSI) overhead and causes big computational burden.

One way to deal with computational complexity in massive MIMO is to use the signal to leakage and noise ratio (SLNR) at the base station (BS) as a performance metric for beamforming and power allocation. Even though signal to interference and noise ratio (SINR) based optimization problems have been thoroughly studied, it is known that they lack simple closed-form solutions due to the complexity and the coupled nature [4]. As a result, many researchers focused their studies on the SLNR which considers the leakage power, i.e. the interference produced from the signal intended for a desired user onto the other users, which decouples the optimization problems at the BSs and so it reduces the complexity. Thus, the SLNR has been successfully applied in MIMO systems for different beamforming and power optimization problems and it is shown to be a reliable metric which improves the system performance. For instance, in [4] and [5] a SLNR based beamforming is applied in a single cell MIMO system and studies on the effect of imperfect CSI on the system performance are conducted. In [6]–[8], the power allocation for a single cell has been investigated assuming an SLNR based beamforming.

In large scale systems with huge numbers of antennas and users, the big CSI overhead becomes critical since the CSI is described through large matrices whose dimensions increase with the number of antennas and users. To overcome this challenge, a two-stage beamforming has been studied [9]–[13] where an outer and an inner beamformers are defined to deal with the statistical and the instantaneous channel changes, respectively. In the first stage, the outer beamformer reduces the channel dimensionality, i.e., the size of the CSI matrices, and so it produces effective channel. During the second stage, the inner beamformer utilizes conventional linear beamforming using only the low-dimensional effective channel. Nevertheless, the two-stage beamforming has the strong assumption that there are user groups and all users in one group have the same channel correlation matrix. Particularly, in the very scarcely researched multicellular cases, this assumption becomes very impractical. For instance, in [12] all users in a cell are closely spaced and have the same channel correlation. Hence, the two-stage beamforming leads to implementation problems like user grouping and, therefore, the performance for more realistic scenarios is still unknown [13].

Therefore, differently from previous works, we study distributed two-stage beamforming in the downlink of a multicellular network where instead of grouping users and assuming that the users in each group share the same statistics, we consider the realistic case that they are randomly distributed in the cells and have independent channel properties. We consider that at each BS, the outer beamformer uses only statistics to reduce the system dimensionality and the inner beamformer performs regularized zero forcing (RZF) on the low-dimensional effective channel. The objective of the outer beamformer at each BS is to reduce the dimensionality so that every processing step after the first stage works with the less computationally demanding effective channels, while at the same time the system performance remains in the same order.

To achieve a fully distributed system, each BS uses SLNR as a performance metric. Unlike the SINR, the SLNR metric decomposes the system parameters such that the optimization problems at each BS become independent of the decisions of the other BSs. The first problem which we face is to compute the SLNR in the case of statistical CSI, since working with statistics is usually a very complex and time-consuming task. Therefore, our first contribution lies in the application of random matrix theory through which we derive a deterministic equivalent of the SLNR in the distributed multicell system which is a very tight approximation in closed-form. Having these approximations, the BS can reliably and fast design the outer beamforming using only statistics. The second contribution of this work consists of the proposed outer beamforming design and power allocation. We propose a low-complex outer beamforming based on block diagonalization which designs a transmission subspace which maximizes the deterministic equivalent of the minimum SLNR in the cell while reducing the dimensionality of the initial transmission space. Moreover, to elaborate on user fairness, we apply power allocation which maximizes the minimum effective SLNR of the users in the cell using only the low-dimensional effective channels. Luckily, the max min SLNR problem can be reformulated to a convex problem and efficiently solved by any convex solver. Simulation results confirm that the proposed two-stage beamforming reduces the system dimensionality while it keeps the performance in the sense of throughput and user fairness very close to the performance of the single stage RZF. Moreover, combining the two-stage beamforming with the power allocation results also in a system with reduced dimensionality and preserved throughput and fairness performance compared to the trivial RZF with the same power allocation approach.

The rest of the paper is organized as follows. In Section II, we present the system model and in Section III the two-stage beamforming. Section IV shows the deterministic equivalent of the SLNR and the design of the outer beamforming. Section V shows the proposed power allocation. Section VI presents simulation results and Section VII concludes of the work.

Notations - Lower case and upper case boldface letters denote vectors and matrices, respectively. The ith entry of vector $\mathbf{x}$ is denoted by $[\mathbf{x}]_i$ and the (i,j)th entry of the matrix $\mathbf{X}$ as $[\mathbf{X}]_{i,j}$. Hermitian transpose and trace of a matrix are denoted as $(\cdot)^{\mathrm{H}}$ and $\mathrm{tr}(\cdot)$, respectively. $\mathbf{I}_N$ stands for identity matrix of size $N \times N$. Euclidean norm of vector $\mathbf{x}$ is denoted as $||\mathbf{x}||$ and spectral norm of matrix $\mathbf{X}$ as $||\mathbf{X}||$. The cardinality of set $\mathcal{A}$ is $|\mathcal{A}|$.

II. System Model

We consider the downlink of a multicellular network with L BSs, each equipped with M_l antennas. Every BS serves a set $\mathcal{K}_l$ of single-antenna users simultaneously, so that the overall number of users in the network is $K = \sum_{l=1}^{L} K_l$ where $K_l = |\mathcal{K}_l|$. To denote the serving BS of user k, $k = 1, \ldots, K$, we use the index $l_k \in \{1, \ldots, L\}$. The data symbol for user k is denoted by s_{k,l_k} and it is modeled as zero mean Gaussian process with variance one, i.e. $s_{k,l_k} \sim \mathcal{CN}(0,1)$.

We consider a typical channel model, known as one-ring channel model [14], where the spatial correlation assumes an elevated BS without surrounding objects and users which have local scattering. The channel between user k and BS l is

$$\mathbf{h}_{k,l} = \mathbf{\Theta}_{k,l}^{1/2} \mathbf{z}_{k,l} \text{ for } l = 1, \ldots, L, \tag{1}$$

with $\mathbf{\Theta}_{k,l}$ the second-order channel statistics which change slowly over time and which can be expressed as a function of the long-term path loss $a_{k,l}$ and the channel correlation $\bar{\mathbf{\Theta}}_{k,l} \in \mathbb{C}^{M_l \times M_l}$, i.e., $\mathbf{\Theta}_{k,l} = a_{k,l} \bar{\mathbf{\Theta}}_{k,l}$. The fast channel fluctuations which change several orders faster than the channel statistics are modeled by $\mathbf{z}_{k,l} \in \mathbb{C}^{M_l \times 1}$ as a random process with identically and independently distributed (i.i.d.) entries taken from a zero mean Gaussian distribution with unit variance, i.e. $\mathbf{z}_{k,l} \sim \mathcal{CN}(0, \mathbf{I}_{M_l})$. Hence, the received signal at user k is

$$\begin{aligned} y_k = {} & \mathbf{h}_{k,l_k}^{\mathrm{H}} \sqrt{p_{k,l_k}} \mathbf{v}_{k,l_k} s_{k,l_k} + \sum_{i \in \mathcal{K}_{l_k}, i \neq k} \mathbf{h}_{k,l_k}^{\mathrm{H}} \sqrt{p_{i,l_k}} \mathbf{v}_{i,l_k} s_{i,l_k} \\ & + \sum_{l=1, l \neq l_k}^{L} \sum_{j \in \mathcal{K}_l} \mathbf{h}_{k,l}^{\mathrm{H}} \sqrt{p_{j,l}} \mathbf{v}_{j,l} s_{j,l} + n_k \end{aligned} \tag{2}$$

with $\mathbf{v}_{k,l} \in \mathbb{C}^{M_l \times 1}$ the beamforming vector at BS l for user k, $p_{k,l}$ the allocated power and n_k the white Gaussian noise at user k with zero mean and variance $\sigma^2 = 1$. From (2), one can see that the received signal y_k in the multicellular network consists of four terms, i.e., useful signal, intra-cell interference, inter-cell interference and noise.

III. Two-Stage Beamforming

In this section, we present the proposed two-stage beamforming which is designed in a fully decentralized manner at each BS where we consider that each BS has statistical CSI from the users in its vicinity, i.e., served users and users from other cells with strong links, and instantaneous effective CSI only from the users which it serves. The beamforming at BS l is denoted as $\mathbf{V}_l = [\mathbf{v}_{i,l}]_{i \in \mathcal{K}_l} \in \mathbb{C}^{M_l \times K_l}$ and we decompose it into two beamformers, i.e. $\mathbf{V}_l = \mathbf{F}_l \mathbf{G}_l$.

The outer beamforming $\mathbf{F}_l \in \mathbb{C}^{M_l \times N_l}$ deals with the transmission subspace dimensionality N_l and the inter-cell interference. It is designed in the first stage of the beamforming design and uses only the channel statistics $\mathbf{\Theta}_{k,l}$. Therefore, it is updated only once for a long time period. Its main purpose is to design a transmission subspace for BS l which has the size N_l such that $N_l \leq M_l$. This produces an effective system of dimension N_l where the effective channel between user k and BS l is the projection of the original channel onto the transmission subspace spanned by $\mathbf{F}_l$, i.e. $\tilde{\mathbf{h}}_{k,l} = \mathbf{F}_l^{\mathrm{H}}\mathbf{h}_{k,l} \in \mathbb{C}^{N_l \times 1}$. Moreover, because working with statistics is a complex task, we derive deterministic equivalent of the SLNR which has closed-form and which is used for the design of the outer beamformers, see Section IV for details.

In the second stage, the inner beamformer $\mathbf{G}_l = [\mathbf{g}_{i,l}]_{i\in\mathcal{K}_l} = [\chi_{i,l}\hat{\mathbf{g}}_{i,l}]_{i\in\mathcal{K}_l}$ with $\chi_{i,l} = 1/||\hat{\mathbf{g}}_{i,l}|| \in \mathbb{C}^{N_l\times K_l}$ uses the fast changing effective channels of the users served by BS l, i.e. $\tilde{\mathbf{h}}_{i,l}$ for $i \in \mathcal{K}_l$, to combat intra-cell interference. Due to the fast variation of the instantaneous CSI, the updates of $\mathbf{G}_l$ are few orders more frequent as compared to the updates of $\mathbf{F}_l$. For $\tilde{\mathbf{H}}_l = [\tilde{\mathbf{h}}_{i,l}]^{\mathrm{H}}_{i\in\mathcal{K}_l} \in \mathbb{C}^{K_l\times N_l}$ the effective channel at BS l and $\hat{\mathbf{g}}_{i,l}$ is the precoding vector for user $i \in \mathcal{K}_l$, the inner beamformer which performs RZF [15] is defined as

$$\hat{\mathbf{G}}_l = [\hat{\mathbf{g}}_{i,l}]_{i\in\mathcal{K}_l} = \left(\tilde{\mathbf{H}}_l^{\mathrm{H}}\tilde{\mathbf{H}}_l + M_l\alpha_l\mathbf{I}_{M_l}\right)^{-1}\tilde{\mathbf{H}}_l^{\mathrm{H}}, \qquad (3)$$

with α_l regularization parameter to control the interference in the cell. We have set $\alpha_l = (K_l\sigma^2)/(P_l M_l)$ to maximize the SINR in the cell [16] where P_l is the power budget at BS l.

IV. Dimensionality Reduction Using Deterministic Equivalents

The SLNR for user k in the multicellular network is

$$\beta_k = \frac{p_{k,l_k}|\mathbf{h}^{\mathrm{H}}_{k,l_k}\mathbf{F}_{l_k}\mathbf{g}_{k,l_k}|^2}{\sum_{i=1,i\neq k}^{K} p_{k,l_k}|\mathbf{h}^{\mathrm{H}}_{i,l_k}\mathbf{F}_{l_k}\mathbf{g}_{k,l_k}|^2 + \sigma^2} \qquad (4)$$

where p_{k,l_k} is the power allocated at BS l_k for user k. The nominator in (4) describes the useful power at user k while the left term in the denominator describes the leakage power which is the interference produced from transmission to user k to the other users in the system.

To design the outer beamforming, we use the SLNR as a performance metric since it decouples the optimization problems at the BSs as it can be seen in (4) where all terms depend only on the design parameters of the serving BS l_k and, hence, (4) is fully independent on the design decisions of any other BS. However, as mentioned above, for the design of the outer beamforming, the BS uses only statistical CSI which means that the exact values of the channels $\mathbf{h}^{\mathrm{H}}_{t,l_k}$ and the RZF vectors $\mathbf{g}_{t,l_k}$ for $t \in \mathcal{K}_{l_k}$ are unknown. Therefore, to make the design in the first stage possible using only the statistical channel knowledge $\mathbf{\Theta}_{i,l}$, we derive a deterministic equivalent which tightly approximates the SLNR value.

A. Deterministic Equivalents

The deterministic equivalent is a tool from random matrix theory and it provides asymptotic expressions of functionals with random matrices whose dimensions increase to infinity [17]. It is already proven that the asymptotic expressions can be used in certain MIMO systems as very accurate approximation of diverse parameters, e.g. [18]–[21]. Moreover, the deterministic equivalents have closed-form which saves computational resources compared to conventional averaging techniques.

For the first-stage beamforming, the CSI is a random matrix with known statistics whose dimensions depend on the number of users and antennas. We denote the deterministic equivalent of a functional x by $\mathring{x}$ where $x - \mathring{x} \xrightarrow{a.s.} 0$. Here, "$\xrightarrow{a.s.}$" presents an almost sure convergence as the dimensions M_l and K_l grow to infinity with fixed ratio between them, i.e. $M_l, K_l \to \infty$ with ratio $c_l = M_l/K_l$ such that $0 < \liminf_{M_l,K_l} c_l \leq \limsup_{M_l,K_l} c_l < \infty$ for $\forall l$.

The derived deterministic equivalent of the SLNR at user k assuming an equal power allocation takes the form

$$\mathring{\beta}_k = \frac{[\mathbf{e}_{l_k}]_k}{1 + \frac{1}{M_{l_k}[\mathbf{e}_{l_k}]_k}\sum_{l=1,l\neq l_k}^{L}\sum_{j\in\mathcal{K}_l}[\mathbf{d}_{k,l_k}]_j}. \qquad (5)$$

To calculate $\mathring{\beta}_k$, we need to compute the set of equations:

$$\mathbf{e}_l = \left[\frac{1}{M_l}\mathrm{tr}(\mathbf{F}_l^{\mathrm{H}}\mathbf{\Theta}_{i,l}\mathbf{F}_l\mathbf{T}_l)\right]_{i\in\mathcal{K}_l} \in \mathbb{C}^{K_l\times 1}, \qquad (6a)$$

$$\mathbf{T}_l = \left(\frac{1}{M_l}\sum_{j\in\mathcal{K}_l}\frac{\mathbf{F}_l^{\mathrm{H}}\mathbf{\Theta}_{j,l}\mathbf{F}_l}{(1+[\mathbf{e}_l]_j)} + \alpha_l\mathbf{I}_{M_l}\right)^{-1}, \qquad (6b)$$

$$[\mathbf{J}_l]_{m,n} = \frac{\mathrm{tr}(\mathbf{F}_l^{\mathrm{H}}\mathbf{\Theta}_{m,l}\mathbf{F}_l\mathbf{T}_l\mathbf{F}_l^{\mathrm{H}}\mathbf{\Theta}_{n,l}\mathbf{F}_l\mathbf{T}_l)}{M_l^2(1+[\mathbf{e}_l]_n)^2} \text{ for } m,n \in \mathcal{K}_l, \qquad (6c)$$

$$\mathbf{c}_{k,l} = (\mathbf{I}_{K_l} - \mathbf{J}_l)^{-1}\mathbf{o}_{k,l} \in \mathbb{C}^{K_l\times 1}, \qquad (6d)$$

$$\mathbf{o}_{k,l} = \left[\frac{1}{M_l}\mathrm{tr}(\mathbf{F}_l^{\mathrm{H}}\mathbf{\Theta}_{m,l}\mathbf{F}_l\mathbf{T}_l\mathbf{F}_l^{\mathrm{H}}\mathbf{\Theta}_{k,l}\mathbf{F}_l\mathbf{T}_l)\right]_{m\in\mathcal{K}_l} \in \mathbb{C}^{K_l\times 1} \qquad (6e)$$

$$\mathbf{b}_{k,l_k} = \left[\frac{1}{M_{l_k}}\mathrm{tr}(\mathbf{F}_{l_k}^{\mathrm{H}}\mathbf{\Theta}_{j,l_k}\mathbf{F}_{l_k}\mathbf{T}_{l_k}\mathbf{F}_{l_k}^{\mathrm{H}}\mathbf{\Theta}_{k,l_k}\mathbf{F}_{l_k}\mathbf{T}_{l_k})\right]_{j\in\mathcal{K}_l}, \qquad (6f)$$

$$\mathbf{M}_{k,l} = \mathbf{T}_l\left(\sum_{m\in\mathcal{K}_l}\frac{\mathbf{F}_l^{\mathrm{H}}\mathbf{\Theta}_{m,l}\mathbf{F}_l[\mathbf{c}_{k,l}]_m}{(1+[\mathbf{e}_l]_m)^2}\right)\mathbf{T}_l/M_l, \qquad (6g)$$

$$\mathbf{d}_{k,l_k} = \left[\mathrm{tr}(\mathbf{F}_{l_k}^{\mathrm{H}}\mathbf{\Theta}_{j,l_k}\mathbf{F}_{l_k}\mathbf{M}_{k,l_k})/M_{l_k} + [\mathbf{b}_{k,l_k}]_j\right]_{j\in\mathcal{K}_l} \qquad (6h)$$

with $j \in \mathcal{K}_l$ for $l = 1, \ldots, L$ and $l \neq l_k$ and, hence, the size of $\mathbf{b}_{k,l_k}$ and $\mathbf{d}_{k,l}$ is $(K - K_{l_k}) \times 1$. *Proof:* See Appendix A.

Every BS can reliably approximate the actual SLNR β_k by $\mathring{\beta}_k$ using only the statistical channel knowledge $\mathbf{\Theta}_{k,l}$. Moreover, because the derivations are described as functions of the outer beamformers $\mathbf{F}_l$, each BS can design its own outer beamforming using the SLNR metric.

B. Outer Beamforming Design

Every BS l designs its own outer beamforming $\mathbf{F}_l$ in a fully distributed manner using $\mathring{\beta}_k$ with $k \in \mathcal{K}_l$. The objective of the outer beamforming is to maximize the minimum $\mathring{\beta}_k$ in the cell and to design a transmission subspace with dimensions N_l such that $K_l \leq N_l \leq M_l$. To achieve this, we propose a low-complex non-iterative approach based on block diagonalization. The basic idea is to find a subspace within the high-dimensional transmission space which introduces only small interference and at the same time has strong contributions for the useful signals. To find this transmission subspace, we use simple linear algebra operations such as singular value decomposition and matrix multiplication. Once a subspace is defined, the minimum $\mathring{\beta}_k$ in the cell is calculated and used as a metric for the performance of the proposed subspace.

To design the outer beamformer $\mathbf{F}_b$ at BS b, we first define a subspace $\mathbf{B}_b^i = \left[\mathbf{F}_b^{\mathrm{H}} \mathbf{\Theta}_{j,b} \mathbf{F}_b\right]_{j \in \{\mathcal{K}_l : l=1,\ldots,L \text{ and } l \neq b\}}$ of size $M_b \times M_b \sum_{l=1, l\neq b}^{L} K_l$ which represents the produced interference from BS b and denote its left-eigenspace spanned by the N_b weakest eigenmodes as $\hat{\mathbf{E}}_b$. Next, we project the transmission subspace $\mathbf{B}_b^s = \left[\mathbf{F}_b^{\mathrm{H}} \mathbf{\Theta}_{i,b} \mathbf{F}_b\right]_{i \in \mathcal{K}_b}$ onto $\hat{\mathbf{E}}_b$ and obtain a matrix from which we take only the N_b strongest eigenmodes from its left-eigenspace and denote it as $\hat{\mathbf{U}}_b$. The outer beamforming is then designed as $\mathbf{F}_b = \hat{\mathbf{E}}_b \hat{\mathbf{U}}_b$ and it defines an N_b-dimensional subspace within the original M_l-dimensional transmission space which considers only the N_b strongest transmission modes producing the least interference.

Additionally, to achieve a good performance, we do not choose the dimension N_b in advance, but let the algorithm find the N_b which maximizes the minimum $\mathring{\beta}_k$ with respect to all $k \in \mathcal{K}_l$. However, to reduce the computational complexity, we let the algorithm to design the interference producing subspace $\mathbf{B}_b^i$ considering only the half of the users to which it leaks interfering power and so its size reduces to $M_b \times M_b (K - K_l)/2$. Moreover, we also examine the system performance when the algorithm searches only from a fixed set of possible dimensions instead of considering all $(M_l - K_l)$ possible dimensions and so we decrease further the computational demands on the outer beamforming. Surprisingly, these additional modifications of the algorithm not only speed up the design, but also preserve the performance in terms of the achieved throughout and fairness.

V. Power Allocation

The outer beamforming produces effective channels with the dimensions $N_l \times 1$ and all further possessing steps work only within this effective domain. Therefore, to improve the fairness and to present the performance of the system for additional processing after applying the proposed two-stage beamforming, we perform power allocation.

To allocate the power, BS l uses the effective SLNR $\tilde{\beta}_k$ based on the fast changing effective channels of its own users

$$\tilde{\beta}_k = \frac{p_{k,l_k} |\tilde{\mathbf{h}}_{k,l_k}^{\mathrm{H}} \mathbf{g}_{k,l_k}|^2}{\sum_{i \in \mathcal{K}_{l_k}, i \neq k} p_{k,l_k} |\tilde{\mathbf{h}}_{i,l_k}^{\mathrm{H}} \mathbf{g}_{k,l_k}|^2 + \sigma^2}. \quad (7)$$

We define a max min power optimization problem where the minimum SLNR $\tilde{\beta}_k$ is maximized with respect to the allocated power such that the power budget at the BS is satisfied, i.e.,

$$\begin{aligned} &\underset{\{p_{k,l} \in \mathbb{R}\}_{k \in \mathcal{K}_l}}{\arg\max} \ \min_{k \in \mathcal{K}_l} \tilde{\beta}_k \\ &\text{s.t.}: \sum_{k \in \mathcal{K}_l} p_{k,l} ||\mathbf{g}_{k,l}|| \leq P_l. \end{aligned} \quad (8)$$

In an analogous manner to [22] where a single-cell power optimization for a multigroup multicast system is considered, one can show that the optimization problem in (8) can be reformulated to a convex problem with an auxiliary positive real variable t and it can be efficiently solved by a standard convex solver, i.e.

$$\begin{aligned} &\min_{\{p_{k,l} \in \mathbb{R}\}_{k \in \mathcal{K}_l}, t \in \mathbb{R}} t^{-1} \\ &\text{s.t.}: \sum_{i \in \mathcal{K}_l, i \neq k} \frac{|\tilde{\mathbf{h}}_{k,l}^{\mathrm{H}} \mathbf{g}_{k,l}|^2}{|\tilde{\mathbf{h}}_{i,l}^{\mathrm{H}} \mathbf{g}_{k,l}|^2} t + \frac{\sigma^2}{|\tilde{\mathbf{h}}_{i,l}^{\mathrm{H}} \mathbf{g}_{k,l}|^2} \frac{t}{p_{k,l}} \leq 1 \quad \text{for } \forall k \in \mathcal{K}_l \\ &\sum_{k \in \mathcal{K}_l} p_{k,l} \leq P_l \\ &p_{k,l} \geq 0 \quad \text{for } \forall k \in \mathcal{K}_l \\ &t \geq 0 \end{aligned} \quad (9)$$

VI. Simulations Results

A. General Setup

In the following simulations, we have considered a system of $L = 9$ or $L = 16$ BSs equipped with $M_l = 64$ antennas which apply the proposed two-stage beamforming and power allocation in a fully distributed way without any coordination between the BSs. In each cell, there are $K_l = 4$ users to be served which are randomly located following a uniform distribution. The cells have a typical hexagonal shape and radius $R_{cell} = 50$ m. The long term path loss between user k and BS l is modeled as $a_{k,l} = 1/(1 + (D_{k,l}/D_0)^{-\gamma_{loss}})$ where $D_{k,l}$ is the distance between the user and the BS, $D_0 = 5$ m is a reference distance and $\gamma_{loss} = 3$ is the path loss exponent.

We assume a uniform linear array at each BS with channel correlation $\bar{\mathbf{\Theta}}_{k,l}$ modeled according to the discrete uniform distribution [23] with random angular spread $\Delta_{k,l}$ of scatterers around the kth user. We have simulated the system for different scattering environments, i.e. for $\Delta_{k,l} \sim \mathcal{U}(1°, 30°)$ and for rich scattering with $\Delta_{k,l} \sim \mathcal{U}(1°, 60°)$.

In the figures below, we have plotted the achieved throughput and the average Jain's fairness index [24] which shows how fairly the BSs serve the users in the system considering the achieved throughput at each user. The worst case is when the index is equal to $1/K_l$ and in the best case it is 1, i.e. all users achieve the same data rate. All simulations are performed for 90 time frames with constant second order statistics and with 100 random channel realizations within each frame. The results are illustrated as functions of the signal to noise ratio (SNR) which is denoted by ρ and defined as the ratio of the power budget at a BS over the noise power.

B. Dimensionality Reduction and System Performance

In the following, *I_RZF* indicates conventional distributed RZF and *F_RZF* the proposed two-stage beamforming. Both *I_RZF* and *F_RZF* have an equal power policy, i.e. $p_{k,l_k} = P_{l_k}/K_{l_k}$ for $\forall k \in \mathcal{K}_{l_k}$. The *I_RZF_p* is the conventional distributed RZF with power allocation and *F_RZF_p* the proposed two-stage beamforming with power allocation, where both use the power control according to Section V. *I_RZF* and *I_RZF_p* work with the original high-dimensional CSI which incurs high computational complexity due to the work with huge matrices and we use them as references for throughput and fairness.

In Fig. 1 are the results for $L = 9$ and $\Delta_{k,l} \sim \mathcal{U}(1°, 30°)$. From the first two subplots, we see that the conventional one-stage and the proposed two-stage approaches achieve very similar sum rate and fairness. Moreover, with the increase of SNR the proposed two-stage beamforming achieves even slightly higher sum rate because by reducing the dimensionality, the algorithm restricts the transmission space of each BS to a subspace such that the SLNR is maximized. In other words, in the high SNR regime, the leakage towards neighboring BSs plays a bigger role in the achieved performance, however, the proposed approach defines a transmission subspace which has cut the dimensions introducing a lot of leakage in the system and, so, the data rate increases. At the same time, this reduction of dimensions produces a low-dimensional equivalent system which is computationally easier to manage. Regarding the fairness, we observe that for *I_RZF* and *F_RZF*, the fairness increases with the SNR because for negligible noise, distributing the power equally among the users results in similar data rate at each user which also means that the BS serves them almost equally, hence, fairness improves. The fairness for the approaches with the proposed power allocation is very high and it only decreases slightly at very high SNR because our power allocation uses only local effective CSI without considering the neighbors. However, due to the high interference towards the neighbors in the high SNR regime, not considering the leakage reduces the system performance.

The third subplot of Fig. 1 shows the selected dimensions N_l chosen by the proposed two-stage beamforming. This box plot is generated with the standard Matlab function "boxplot" and it presents the median, the 25th and the 75th percentiles as well as the outliers. The whisker is approximately $+/-2.7$ of the standard deviation of the data (for normally distributed data this is 99.3%). We have observed in this and in all other simulations that the BSs start to assign less dimensions for SNR values above 20 dB. This is so because at very low SNR, the produced interference is not that prominent, hence, the BSs assign bigger transmission subspaces. When the SNR is higher, the interference starts to affect more the system performance and the outer beamformer restricts the transmission subspace to lower dimensionality. In the particular example in Fig. 1, we observe that for $\rho = 20$, dB the outer beamformers decide to reduce the dimensionality in about 8% of the cases and for $\rho = 35$ dB in about 72% of the cases.

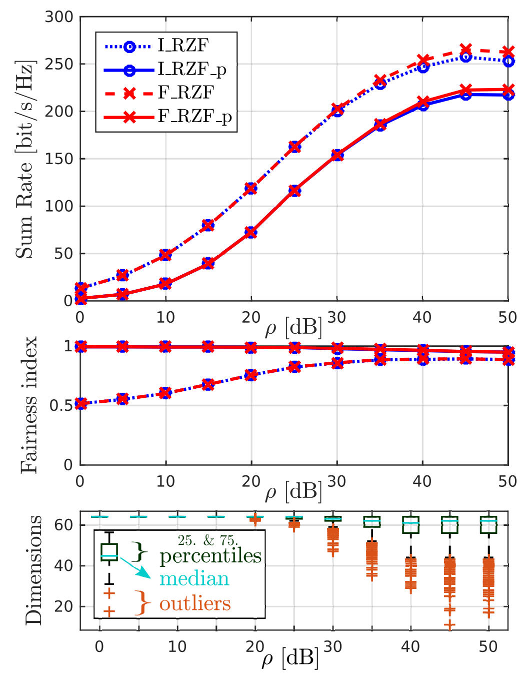

Fig. 1: Performance for $L = 9$ and $\Delta_{k,l} \sim \mathcal{U}(1°, 30°)$

Interestingly, when we restrict the first stage to consider only a small set of possible dimensions in order to reduce the computational time, we achieve very similar performance. The proposed approach with fixed dimensions achieves very similar sum rate and fairness as the conventional distributed approach and, for high SNR, it again slightly outperforms in sum rate. By fixing the possible dimensions, the approach assigns higher dimensions more often than before, however, the trends of the dimensionality reduction remain as in the case when we do not fix them, namely, a significant reduction happens for SNR above 20 dB and with the increase of SNR, the approach reduces the dimensionality even more. Therefore, for the remaining results, without using any statistics, we have chosen the set of possible dimensions to be $\{24, 32, 42, 48, 56, 60, 64\}$. Note that the decision of what number and which particular dimensions can be chosen will define the dimensionality of the equivalent system and the computational speed.

In Fig. 2 we have plotted the results for $L = 9$, $L = 16$, $\Delta_{k,l} \sim \mathcal{U}(1°, 60°)$ and the fixed set of possible dimensions introduced above. Here, we observe that the proposed two-stage beamforming with and without power allocation achieves again very close performance to the conventional approaches while the dimensionality of the effective system is reduced even though only seven different dimensions were allowed. For example, at $\rho = 35$ dB, a reduction happens in 46% of the simulations with $L = 9$ and slightly more than 52% by $L = 16$. Additionally, the scenario with increased number of BSs, which means also more interference in the system, appears to gain even more in data rate thanks to the subspace reduction.

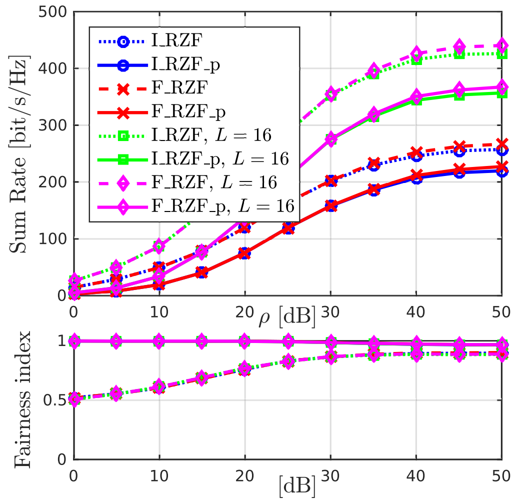

Fig. 2: Performance for $L = 9$, $L = 16$, $\Delta_{k,l} \sim \mathcal{U}(1°, 60°)$ and fixed set of possible dimensions

VII. Conclusions

We propose two-stage beamforming in a distributed multiuser massive MIMO system where in the first stage, an outer beamforming uses only statistics and reduces the system dimensionality while in the second stage, an inner beamformer works with the instantaneous low-dimensional effective CSI to combat interference. We first derive deterministic equivalent of the SLNR which requires only statistical CSI and provides very accurate approximations in easy to compute closed-form. Secondly, we propose a low-complex outer beamforming design using the deterministic equivalents and a power allocation which improves the fairness by optimizing the max min effective SLNR which is shown to be convex problem. Simulation results confirm that the proposed beamforming as well as the proposed beamforming combined with power allocation reduce the system dimensionality while the system performance in terms of achieved throughput and fairness is preserved.

Acknowledgment

This work has been performed in the context of DFG funded CRC 1053 MAKI and supported by DAAD with funds from the Federal Ministry of Education and Research (BMBF).

Appendix A

In the following, we provide a proof of (6) which needs to be solved in order to define the deterministic equivalent of the SLNR $\mathring{\beta}_k$ of (5). We use the following lemmas and theorems:

L1: Matrix inversion lemma, Eq. (2.2) in [25]

L2: Derivative of matrix inverse, Eq. (59) in [26]

L3: Trace lemma, Lemma 14.2 in [17]

L4: Rank-one perturbation lemma, Lemma 2.1 in [27]

T1: Theorem 1 in [20]

T2: Dominated convergence (Theorem 16.4 in [28])

as well as the assumptions:

Assumption 1: We introduce a random variable which is i.i.d. with zero mean and variance $1/M_l$, with finite eighth order moment and define it as $\bar{\mathbf{z}}_{k,l} = \mathbf{z}_{k,l}/\sqrt{M_l}$.

Assumption 2: The effective channel correlation $\tilde{\mathbf{\Theta}}_{k,l} = \mathbf{F}_l^{\mathrm{H}}\mathbf{\Theta}_{k,l}\mathbf{F}_l$ is deterministic with uniformly bounded spectral norm, i.e., $\lim\sup_{M_l\to\infty}\sup_{1\le k\le K}||\tilde{\mathbf{\Theta}}_{k,l}|| < \infty$ for $\forall l$.

Assumption 3: Every BS has an equal power distribution, i.e. $p_{k,l_k} = P_{l_k}/K_{l_k}$.

The first two assumptions are required in order to apply **T1** while the third one is done to simplify the expression since assuming equal power among the users provides significantly simpler expressions and greatly reduced computational complexity for the outer beamforming which uses these derivations.

We define useful power $S_k = |\mathbf{h}_{k,l_k}^{\mathrm{H}}\mathbf{F}_{l_k}\mathbf{g}_{k,l_k}|^2$, power leakage towards the users in its own cell $L_k^{ra} = \sum_{i\in\mathcal{K}_{l_k}, i\neq k}|\tilde{\mathbf{h}}_{i,l_k}^{\mathrm{H}}\mathbf{g}_{k,l_k}|^2$ and towards the users in the neighboring cells $L_k^{er} = \sum_{l=1,l\neq l_k}^{L}\sum_{j\in\mathcal{K}_l}|\tilde{\mathbf{h}}_{j,l_k}^{\mathrm{H}}\mathbf{g}_{k,l_k}|^2$. Hence, (4) becomes

$$\beta_k = \frac{S_k}{L_k^{ra} + L_k^{er} + \sigma^2/p_{k,l_k}} \tag{10}$$

The deterministic equivalent of S_k is $\mathring{S}_k = [\mathbf{e}_{l_k}]_k^2/(1 + [\mathbf{e}_{l_k}]_k)^2$, derived in one of our previous works, see $\mathring{S}_k$ in [18], where $[\mathbf{e}_l]_j$ is defined through $\mathbf{T}_l$ as in (6b) and (6a). The denominator of (10) can be expressed as $\nu_1 + L_k^{er}$ where, similar to [29] we combine the leakage in the cell with the noise power, to simplify the expressions, hence, $\nu_1 = \mathbf{g}_{k,l_k}^{\mathrm{H}}\left[\sum_{i\in\mathcal{K}_{l_k}, i\neq k}\tilde{\mathbf{h}}_{i,l_k}\tilde{\mathbf{h}}_{i,l_k}^{\mathrm{H}} + \sigma^2/p_{k,l_k}\mathbf{I}_{M_{l_k}}\right]\mathbf{g}_{k,l_k}$ and $L_k^{er} = \mathbf{g}_{k,l_k}^{\mathrm{H}}\left[\sum_{l=1,l\neq l_k}^{L}\sum_{j\in\mathcal{K}_l}\mathbf{h}_{j,l_k}^{\mathrm{H}}\mathbf{F}_{l_k}\right]\mathbf{g}_{k,l_k}$.

Having $\tilde{\mathbf{C}}_l = \tilde{\mathbf{\Gamma}}_l + \alpha_l\mathbf{I}_{M_l}$ with $\tilde{\mathbf{\Gamma}}_l = \frac{1}{M_l}\tilde{\mathbf{H}}_l^{\mathrm{H}}\tilde{\mathbf{H}}_l$, ν_1 becomes $\nu_1 = \frac{1}{M_{l_k}}\tilde{\mathbf{h}}_{k,l_k}^{\mathrm{H}}\tilde{\mathbf{C}}_{l_k}^{-1}\left[\frac{1}{M_{l_k}}\sum_{i\in\mathcal{K}_{l_k}, i\neq k}\tilde{\mathbf{h}}_{i,l_k}\tilde{\mathbf{h}}_{i,l_k}^{\mathrm{H}} + \frac{1}{M_{l_k}}\sigma^2/p_{k,l_k}\mathbf{I}_{M_{l_k}}\right]\tilde{\mathbf{C}}_{l_k}^{-1}\tilde{\mathbf{h}}_{k,l_k}$.

Then, we define $\tilde{\mathbf{\Gamma}}_{[p],l} = \frac{1}{M_l}\tilde{\mathbf{H}}_{[p],l}^{\mathrm{H}}\tilde{\mathbf{H}}_{[p],l}$ with $\tilde{\mathbf{H}}_{[p],l}$ the matrix $\tilde{\mathbf{H}}_l$ without the pth row and apply **L1** leading to $\nu_1 = (\bar{\mathbf{z}}_{k,l_k}^{\mathrm{H}}\tilde{\mathbf{\Theta}}_{k,l_k}^{\mathrm{H}/2}\tilde{\mathbf{C}}_{[k],l_k}^{-1}\tilde{\mathbf{\Theta}}_{k,l_k}^{1/2}\bar{\mathbf{z}}_{k,l_k})/((1 + \bar{\mathbf{z}}_{k,l_k}^{\mathrm{H}}\tilde{\mathbf{\Theta}}_{k,l_k}^{\mathrm{H}/2}\tilde{\mathbf{C}}_{[k],l_k}^{-1}\tilde{\mathbf{\Theta}}_{k,l_k}^{1/2}\bar{\mathbf{z}}_{k,l_k})^2)$. In an analogous manner to the proof of $\mathring{S}_k$, after applying **L3**, **L4** and **T1**, we obtain the deterministic equivalent of $(L_k^{ra} + \sigma^2/p_{k,l_k})$, i.e. $\mathring{\nu}_1 = [\mathbf{e}_{l_k}]_k/(1 + [\mathbf{e}_{l_k}]_k)^2$.

For L_k^{er}, we first apply **L1** and obtain $L_k^{er} = (\bar{\mathbf{z}}_{k,l_k}^{\mathrm{H}}\tilde{\mathbf{\Theta}}_{k,l_k}^{\mathrm{H}/2}\tilde{\mathbf{C}}_{[k],l_k}^{-1}\Big(\sum_{l=1,l\neq l_k}^{L}\sum_{j\in\mathcal{K}_l}\tilde{\mathbf{h}}_{j,l_k}\tilde{\mathbf{h}}_{j,l_k}^{\mathrm{H}}\Big)\tilde{\mathbf{C}}_{[k],l_k}^{-1}\tilde{\mathbf{\Theta}}_{k,l_k}^{1/2}\bar{\mathbf{z}}_{k,l_k})$ $(M_{l_k}(1 + \bar{\mathbf{z}}_{k,l_k}^{\mathrm{H}}\tilde{\mathbf{\Theta}}_{k,l_k}^{\mathrm{H}/2}\tilde{\mathbf{C}}_{[k],l_k}^{-1}\tilde{\mathbf{\Theta}}_{k,l_k}^{1/2}\bar{\mathbf{z}}_{k,l_k})^2)$ and so

$$L_k^{er} = L_k^{er\prime}/M_{l_k}(1 + \bar{\mathbf{z}}_{k,l_k}^{\mathrm{H}}\tilde{\mathbf{\Theta}}_{k,l_k}^{\mathrm{H}/2}\tilde{\mathbf{C}}_{[k],l_k}^{-1}\tilde{\mathbf{\Theta}}_{k,l_k}^{1/2}\bar{\mathbf{z}}_{k,l_k})^2. \tag{11}$$

The denominator of (11) is of already known form and its deterministic equivalent is $(1 + [\mathbf{e}_{l_k}]_k)^2$. Therefore, from here on, we focus only on the numerator in (11) which converges to $\frac{1}{M_{l_k}} \sum_{l=1, l \neq l_k}^{L} \sum_{j \in \mathcal{K}_l} \mathrm{tr}\big(\tilde{\mathbf{\Theta}}_{j,l_k} \tilde{\mathbf{C}}_{l_k}^{-1} \tilde{\mathbf{\Theta}}_{k,l_k} \tilde{\mathbf{C}}_{l_k}^{-1}\big)/M_{l_k}$ after applying **L3** where the expression after the summation equals

$$\begin{aligned} &\frac{\partial}{\partial z} \frac{1}{M_{l_k}} \mathrm{tr}\left(\tilde{\mathbf{\Theta}}_{j,l_k} \big(\tilde{\mathbf{\Gamma}}_{l_k} + \alpha_{l_k} \mathbf{I}_{l_k} - z \tilde{\mathbf{\Theta}}_{k,l_k}\big)^{-1} \right) \Bigg|_{z=0} \\ &= \frac{\partial}{\partial z} f_{j,k,l_k} \Bigg|_{z=0}. \end{aligned} \tag{12}$$

After **T1** $f_{j,k,l_k} \xrightarrow{a.s.} [\hat{\mathbf{d}}_{k,l_k}]_j = \mathrm{tr}(\tilde{\mathbf{\Theta}}_{j,l_k} \hat{\mathbf{T}}_{l_k})/M_{l_k}$ with $\hat{\mathbf{T}}_{l_k} = \left(\frac{1}{M_{l_k}} \sum_{t \in \mathcal{K}_{l_k}} \frac{\tilde{\mathbf{\Theta}}_{t,l_k}}{(1+[\hat{\mathbf{d}}_{k,l_k}]_t)} - z\tilde{\mathbf{\Theta}}_{k,l_k} + \alpha_{l_k} \mathbf{I}_{M_{l_k}} \right)^{-1}$.

$[\hat{\mathbf{d}}_{k,l_k}]_{t \in \mathcal{K}_{l_k}} = [\mathbf{e}_{l_k}]_t$ and $\hat{\mathbf{T}}_{l_k} = \mathbf{T}_{l_k}$ for $z = 0$. After **L2** on $\hat{\mathbf{T}}_{l_k}$, we obtain $\hat{\mathbf{T}}'_{l_k} = \mathbf{T}_{l_k} \left(\frac{1}{M_{l_k}} \sum_{t \in \mathcal{K}_{l_k}} \frac{\tilde{\mathbf{\Theta}}_{t,l_k} [\hat{\mathbf{d}}'_{k,l_k}]_t}{(1+[\mathbf{e}_{l_k}]_t)^2} \right)^{-1} \mathbf{T}_{l_k} + \mathbf{T}_{l_k} \tilde{\mathbf{\Theta}}_{k,l_k} \mathbf{T}_{l_k}$.

The last component which needs to be derived is $\hat{\mathbf{d}}'_{k,l_k}$. For simplicity, we define $\mathbf{c}_{k,l} = [\hat{\mathbf{d}}'_{k,l_k}]_{t \in \mathcal{K}_{l_k}}$ with dimension $K_{l_k} \times 1$ as well as $\mathbf{d}_{k,l_k} = [\hat{\mathbf{d}}'_{k,l_k}]_j$ for $j = 1, \ldots, K$ and $j \notin \mathcal{K}_{l_k}$ with dimension $(K - K_{l_k}) \times 1$.

First, we discus the deterministic equivalent $\mathbf{c}_{k,l}$ which, analogically to the proof of (10g) and (10h) in [18], it can be shown to have the form presented in (6d) by making use of the matrix $\mathbf{J}_{l_k}$ as defined in (6c).

The last term to derive is $\mathbf{d}_{k,l_k}$ and it can be expressed as

$$\mathbf{d}_{k,l_k} = \frac{1}{M_{l_k}} \mathrm{tr}(\tilde{\mathbf{\Theta}}_{j,l_k} \mathbf{M}_{k,l_k}) + \frac{1}{M_{l_k}} \mathrm{tr}(\tilde{\mathbf{\Theta}}_{j,l_k} \mathbf{T}_{l_k} \tilde{\mathbf{\Theta}}_{k,l_k} \mathbf{T}_{l_k}) \tag{13}$$

for $j = 1, \ldots, K$ and $j \notin \mathcal{K}_{l_k}$ and for $\mathbf{M}_{k,l_k}$ as defined in (6g). In the expression above, the term after the sum sign is specified in the set of (6) as $\mathbf{b}_{k,l_k}$. Substituting $\mathbf{c}_{k,l_k}$, $\mathbf{M}_{k,l_k}$ and $\mathbf{b}_{k,l_k}$ into (13) results in $\mathbf{d}_{k,l_k}$ and, hence, the deterministic equivalent of the leakage power towards users in neighboring cells becomes

$$\mathring{L}_k^{er} = \frac{1}{M_{l_k}} \frac{1}{(1 + [\mathbf{e}_{l_k}]_k)^2} \sum_{l=1, l \neq l_k}^{L} \sum_{j \in \mathcal{K}_l} [\mathbf{d}_{k,l_k}]_j. \tag{14}$$

Finally, substituting the deterministic equivalents of the different power terms, i.e. $\mathring{S}_k$, $\mathring{\nu}_1$ and $\mathring{L}_k^{er}$, into the SLNR expression, we obtain $\mathring{\beta}_k = \mathring{S}_k / (\mathring{\nu}_1 + \mathring{L}_k^{er})$ as shown in (5).

References

[1] D. Gesbert, M. Shafi, D. Shiu, P. Smith, and A. Naguib, "From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems," *IEEE J. Sel. Areas in Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[2] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity Limits of MIMO Channels," *IEEE J. Sel. Areas in Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[3] D. Gesbert, M. Kountouris, R. Heath, C. Chae, and T. Saelzer, "Shifting the MIMO Paradigm," *IEEE Signal Processing Magazine*, Sept. 2007.

[4] M. Sadek, A. Tarighat, and A. Sayed, "A Leakage-Based Precoding Scheme for Downlink Multi-User MIMO Channels," *IEEE Trans. on Wireless Commun.*, vol. 6, no. 5, pp. 1711–1721, May 2007.

[5] H. Shen, W. Xu, A. Swindlehurst, and C. Zhao, "Transmitter Optimization for Per-Antenna Power Constrained Multi-Antenna Downlinks: An SLNR Maximization Methodology," *IEEE Trans. on Signal Processing*, vol. 64, no. 10, pp. 2712–2725, May 2016.

[6] W. Fang, H. Sun, and L. Yang, "Power Allocation for Maximizing Sum Capacity of Multiuser MIMO Downlink with Transmit Precoding Based on SLNR," *IEEE 73rd Vehicular Technology Conf.*, pp. 1–5, May 2011.

[7] W. Jing, and X. Zhong, "Dynamic Power Allocation Based on SLNR Precoding for Multi-User MIMO Downlink," *4th Int. Conf. on Wireless Commun., Networking and Mobile Computing*, pp. 1–4, Oct. 2008.

[8] J. Wang, X. Wang, Y. Guo, and X. You, "A channel adaptive power allocation scheme based on SLNR precoding for multiuser MIMO systems," *IEEE 72nd Vehicular Technology Conf.*, pp. 1–4, Sept. 2010.

[9] G. Lee, and Y. Sung, "A New Approach to User Scheduling in Massive Multi-User MIMO Broadcast Channels," *IEEE Trans. on Commun.*, vol. 66, no. 4, pp. 1481–1495, April 2018.

[10] D. Kim, G. Lee, and Y. Sung, "Two-Stage Beamformer Design for Massive MIMO Downlink By Trace Quotient Formulation," *IEEE Trans. on Commun.*, vol. 63, no. 6, pp. 2200–2211, Jun. 2015.

[11] Y. Jeon, H. Ku, and N. Lee, "Beamforming design and user grouping for FDD massive MIMO systems," *18th IEEE Int. Workshop Signal Process. Adv. Wireless Commun.*, pp. 439–443, Oct. 2017.

[12] J. Chen, and V. Lau, "Two-tier precoding for FDD multi-cell massive MIMO time-varying interference networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1230–1238, Jun. 2014.

[13] Y. Jeon, and M. Min, "Large System Analysis of Two-Stage Beamforming With Limited Feedback in FDD Massive MIMO Systems," *IEEE Trans. on Vehicular Technology*, vol. 67, no. 6, pp. 4984–4997, Jun. 2018.

[14] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. on Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[15] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit wiener filter for the downlink of TDD DS-CDMA systems," *7th IEEE ISSSTA*, vol. 1, no. 1, pp. 9–13, Sept. 2002.

[16] V. K. Nguyen and J. S. Evans, "Multiuser transmit beamforming via regularized channel inversion: A large system analysis," *IEEE Global Telecommunications Conference*, pp. 1–4, Dec. 2008.

[17] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Commun.*. Cambridge University Press, 2011.

[18] B. Boiadjieva, H. Al-Shatri, and A. Klein, "Hierarchical Beamforming in CRAN Using Random Matrix Theory," *2018 IEEE International Conf. on Commun.*, Jun. 2018.

[19] B. Boiadjieva, H. Al-Shatri, and A. Klein, "Scheduling Strategies for Hierarchical Beamforming in Cloud RAN," *22nd Int. ITG Workshop on Smart Antennas*, Mar. 2018.

[20] 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. on Information Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.

[21] J. Zhang, C. Wen, S. Jin, X. Gao, and K. Wong, "Large system analysis of cooperative multi-cell downlink transmission via regularized channel inversion with imperfect CSIT," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 10, pp. 4801–4813, Oct. 2013.

[22] E. Karipidis, N. Sidiropoulos, and Z. Luo, "Quality of Service and Max-Min Fair Transmit Beamforming to Multiple Cochannel Multicast Groups," *IEEE Trans. on Signal Processing*, vol. 56, no. 3, pp. 1268–1279, Mar. 2008.

[23] R. Ertel, P. Cardieri, K. Sowerby, T. Rappaport, and J. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Communications*, vol. 5, no. 1, pp. 10–22, Feb 1998.

[24] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system.," *IEEE Personal Communications*, vol. 5, no. 1, pp. 10–22, Feb 1998.

[25] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *J. of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.

[26] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. Technical University of Denmark, Nov. 2012.

[27] Z. Bai and J. Silverstein, "On the signal-to-interference-ratio of CDMA systems in wireless communications," *The Annals of Applied Probability*, vol. 17, pp. 81–101, Feb. 2007.

[28] P. Billingsley, *Probability and Measure*, 3rd ed. John Wiley & Sons, Inc., 1995.

[29] S. Park, J. Park, A. Yazdan,and R. Heath, "Optimal User Loading in Massive MIMO Systems With Regularized Zero Forcing Precoding," *IEEE Wireless Commun. Letters*, vol. 6, no. 1, pp. 118–121, Feb. 2017.