Decentralizing the Optimal Multi-cell Beamforming via Large System Analysis

Hossein Asgharimoghaddam, Antti Tölli and Nandana Rajatheva
Centre for Wireless Communications, University of Oulu, P.O. Box 4500, FIN-90014, Oulu, Finland.
{hasghari, atolli, rrajathe}@ee.oulu.fi

***Abstract*—A multi-cell minimum power beamforming problem is considered. It is known that the inter-cell interference (ICI) terms couple the base stations (BS) and inter-cell coordination is required for global optimal solution. The cooperation can be realized by exchange of instantaneous channel state information (CSI) or terms related to the ICI values via a backhaul link. However, the limited backhaul capacity and delay constraints put a limit on achievable performance when the number of antennas and users grow large or when dealing with a fast fading scenario. In this work, we demonstrate that the ICI terms coupling the coordinating BSs can be approximated using the random matrix theory (RMT) tools when the problem dimensions grow large, and that the approximated ICI values depend only on channels statistics, i.e., spatial load and user specific path loss values. This leads to a significant reduction in the information exchange rate among BSs. Furthermore, processing is simplified because with a fixed approximated ICI values the beamforming vectors can be obtained locally at each BS. The proposed solution guarantees the feasibility of the target signal-to-interference-plus-noise ratios (SINR) without any major loss of performance as compared to the optimal centralized design.**

I. Introduction

Coordinated multi-point transmission (CoMP), which is already considered in 3GPP standard, has been studied widely for cellular systems. CoMP allows cooperation and coordination between nodes for delivering services to users which results in greatly improved optimization objective values as compared to the non-coordinated transmission [1]. In order to realize the gains in practice, the coordinating base stations must share some information. The coordination can be divided in two categories; joint transmission and coordinated beamforming/scheduling. Joint transmission from several coordinating nodes, also called as network MIMO [2], ideally provides the largest improvement of the network optimization objective. However, it requires the user data and the channel state information (CSI) to be shared between the coordinated nodes making the practical implementation often difficult, especially when dealing with a large number of users and serving nodes scattered in a large network. As an alternative approach, the coordinated beamforming/scheduling can be considered, where the user data is just available in its serving node and decisions about beamforming/scheduling are made jointly by sharing some CSI information [3]. Maximizing a network optimization objective, subject to some additional quality of service constraints usually results in an iterative distributed solution which requires exchange of some limited information between nodes at each iteration [3]–[8].

Coordinated multi-cell minimum power beamforming approach satisfies a given signal-to-interference-plus-noise ratio (SINR) for all users while minimizing the total transmitted power, which not only fulfills the quality of service for all users, irrespective of their distances, but also minimizes the interference level in general. Authors in [7] solve this problem using the duality in convex optimization theory which is shown to be linked to the concept of uplink-downlink duality. These solutions do not require sharing the user's data; however, the nodes should exchange their local user's instantaneous CSI. Sharing instantaneous CSI between nodes under a delay constraint and limited backhaul capacity becomes a principal problem when dimensions of the problem (number of antennas N and number of users K) grow large or when dealing with a fast fading scenario. Authors in [9] have extended the work in [7] for a large dimension system where random matrix theory (RMT) is utilized to give approximate beamformers at each BS, which rely only on the local CSI and average channel statistics (spatial load, pathloss information) from the other BS channels. However, the target SINR feasibility cannot be guaranteed as the error in approximations is translated into variation in the resulting SINR values.

In [8], [10], [11], an alternative decentralized framework is proposed for the coordinated multi-cell minimum power beamformer design problem. The optimal minimum power beamformers can be obtained locally at each base station (BS) relying on limited backhaul information exchange between BSs. The proposed method is able to guarantee feasible solutions even if the interference information is outdated or incomplete. As alternatives to convex optimisation solutions, iterative fixed-point algorithms based on, for example, Lagrangian duality theory have been developed for solving local optimization problems at each BS [11].

In this work, the focus is on large-scale multiple antenna wireless systems with a large number of low-power antennas co-located at the BS site, often called as *massive MIMO* [12], [13]. One important benefit of such a setting is that some of the analysis can be carried out using tools from random matrix theory (RMT) [13]. It was shown in [12] that with very large imbalance $N >> K$, the processing can be simplified in a way that even matched filter (MF) and zero-forcing (ZF) can be used in an ideal i.i.d. channel for near optimal detection and beamforming [12], [13]. However, in practical multi-cell environments with non-ideal, correlated channels,

Part of this work has been performed in the framework of the FP7 project ICT-317669 METIS, which is partly funded by the European Union.

the use of more complicated precoder design algorithms is justified as the performance gains compared to simple MF or ZF based schemes are still significant. We propose a novel approach for decoupling the subproblems at base stations. Following the same logic as in [8], inter-cell interference (ICI) is considered as the principal coupling parameter among BSs. We use a large dimension approximation for ICI term based on random matrix theory which leads to a distributed beamforming algorithm. This algorithm gives the feasible beamforming vectors at each BS based on the local CSI only, while the coupling ICI terms are based on acquired channel statistics (spatial load, propagation loss values) from the other BS channels. The proposed algorithm benefits from the large dimension simplifications and performance gains due to the limited coordination between nodes which results in almost optimal transmit powers along with significant reduction in the backhaul exchange rate.

II. System Model

A cellular system is considered which consists of N_{B} BSs, each BS has N_{a} transmit antennas. Each user has a single receive antenna. Users allocated to the b^{th} base station are in set $\mathcal{U}_b$. Each user is served by a single base station and the BS that serves user k is denote by b_k. Sets of all users and all BSs are presented by $\mathcal{U}$ and $\mathcal{B}$ respectively. The signal for user k consists of the desired signal, the intracell and the intercell interference which can be presented as follows,

$$\mathbf{y}[k] = \mathbf{h}_{b_k,k}\mathbf{x}_{b,k} + \mathbf{h}_{b_k,k}\sum_{l\neq k\in\mathcal{U}_{b_k}}\mathbf{x}_{b,l} + \sum_{b\neq b_k}\mathbf{h}_{b,k}\sum_{l\in\mathcal{U}_b}\mathbf{x}_{b,l} + \mathbf{n}_k \tag{1}$$

where $\mathbf{n}_k \sim \mathcal{CN}(0, N_0)$ is the noise with power density N_0. $\mathbf{h}_{b,k} \in \mathbb{C}^{1\times N_{\mathrm{a}}}$ represents the channel from the b^{th} BS to k^{th} user. The elements of channels are assumed to be Gaussian distributed and the path-losses $a_{b,k}^2$ are included in the channel vectors, i.e, $\mathbf{h}_{b,k} \sim \mathcal{CN}(0, a_{b,k}^2\mathbf{I}_{\mathrm{N_a}})$. $\mathbf{x}_{b,k} = \mathbf{w}_{b,k}d_k$ is the transmitted vector from the b^{th} BS to k^{th} user, in which d_k is the normalized complex data symbol ($E[|d_k|^2] = 1$) and $\mathbf{w}_{b,k} \in \mathbb{C}^{N_{\mathrm{a}}}$ is the downlink beamforming vector from the b^{th} BS to k^{th} user.

III. Problem Formulation

The optimization problem for achieving the optimal downlink beamformers as proposed by [8] can be presented as

$$\begin{aligned} \underset{\mathbf{w}_{b,k},\epsilon_{b,k}}{\text{minimize}} \quad & \sum_{b\in\mathcal{B}}\sum_{k\in\mathcal{U}_b}\|\mathbf{w}_{b,k}\|^2 \\ \text{subject to} \quad & \Gamma_k \geq \gamma_k \ \forall k\in\mathcal{U}_b, \forall b, \\ & \sum_{l\in\mathcal{U}_b}|\mathbf{h}_{b,k}\mathbf{w}_{b,l}|^2 \leq \epsilon_{b,k}^2, \forall k\notin\mathcal{U}_b, \forall b, \end{aligned} \tag{2}$$

where the intercell interference from b^{th} base station to user k is denoted by $\epsilon_{b,k}^2$, and where

$$\Gamma_k = \frac{|\mathbf{h}_{b_k,k}\mathbf{w}_{b_k,k}|^2}{N_0 + \sum_{l\in\mathcal{U}_{b_k}\setminus k}|\mathbf{h}_{b_k,k}\mathbf{w}_{b_k,l}|^2 + \sum_{b\neq b_k}\epsilon_{b,k}^2} \tag{3}$$

This formulation highlights the role of intercell interference in coupling the beamforming subproblems at base stations. Note that both the SINR and the ICI constraints hold with equality at the optimal solution. The optimization problem defined by (2) can be formulated as a second order cone problem (SOCP) and be solved in a centralized manner by using convex optimization tools [8].

A. Decentralized solutions via optimization decomposition

The centralized problem in (2) is decoupled among BSs as soon as the ICI terms $\epsilon_{b,k}$ are set to fixed values. In [8], the coupling is handled by taking the local copies of the interference terms at each BS and enforcing consistency between them. Then, the consistency constraints become decoupled by applying a standard dual decomposition approach that results a distributed algorithm. The decentralized algorithm can follow the optimal solution in a time correlated scenario by exchanging the ICI terms while the channel realizations change. There are also alternative decentralized solutions based on primal decomposition [10], [11] and alternating direction method of multipliers (ADMM) [14].

B. Solution via uplink-downlink duality

Another approach for solving the optimization problem defined by (2) is based on uplink-downlink duality. Authors in [7] have shown that the problem dual to (2) which gives the optimal uplink power allocation and detection vectors is defined as follows

$$\begin{aligned} \underset{\hat{\mathbf{w}},\lambda}{\text{minimize}} \quad & \sum_{b\in\mathcal{B}}\sum_{k\in\mathcal{U}_b}\lambda_k N_0 \\ \text{subject to} \quad & \frac{\lambda_k|\hat{\mathbf{w}}_{b_k,k}^{\mathrm{H}}\mathbf{h}_{b_k,k}^{\mathrm{H}}|^2}{\sum_{l\neq k}\lambda_l|\hat{\mathbf{w}}_{b_l,l}^{\mathrm{H}}\mathbf{h}_{b_l,k}^{\mathrm{H}}|^2 + \|\hat{\mathbf{w}}_{b_k,k}\|^2} \geq \gamma_k \ \forall k\in\mathcal{U} \end{aligned} \tag{4}$$

The dual uplink power of the k^{th} user is denoted by λ_k that its optimal value can be calculated by a fixed point iteration [7]

$$\lambda_k = \frac{1}{(1+\frac{1}{\gamma_k})\mathbf{h}_{b_k,k}(\mathbf{\Sigma}_{b_k}+\mathbf{I})^{-1}\mathbf{h}_{b_k,k}^{\mathrm{H}}} \tag{5}$$

where

$$\mathbf{\Sigma}_b = \sum_{l\in\mathcal{U}}\lambda_l\mathbf{h}_{b,l}^{\mathrm{H}}\mathbf{h}_{b,l} \tag{6}$$

The dual uplink detection vector $\hat{\mathbf{w}}_{b,k}$ is given by the minimum mean square error receiver at the optimal point [7], i.e,

$$\hat{\mathbf{w}}_{b,k} = (\mathbf{\Sigma}_b + \mathbf{I})^{-1}\mathbf{h}_{b,k}^{\mathrm{H}} \tag{7}$$

A link between the downlink and uplink beamformers is provided by the following equation [7]

$$\mathbf{w}_{b_k,k} = \sqrt{\delta_k}\hat{\mathbf{w}}_{b_k,k} \tag{8}$$

where $\mathbf{w}_{b,k}$ is the downlink beamformer for the k^{th} user and δ_k can be found by the following matrix inversion [7]

$$G_{i,j} = \begin{cases} \frac{1}{\gamma_i}|\hat{\mathbf{w}}_{b_i,i}^{\mathrm{H}}\mathbf{h}_{b_i,i}^{\mathrm{H}}|^2 & i=j \\ -|\hat{\mathbf{w}}_{b_j,j}^{\mathrm{H}}\mathbf{h}_{b_j,i}^{\mathrm{H}}|^2 & i\neq j. \end{cases} \tag{9}$$

Finally,

$$\boldsymbol{\delta} = \mathbf{G}^{-1}\mathbf{1}_{Nu} \tag{10}$$

where $\boldsymbol{\delta}$ is a vector that contains all δ_k values and $\mathbf{1}_{Nu}$ is a $Nu \times 1$ vector with all elements equal to one.

The above set of equations defines an algorithm which gives the optimal power allocation and beamformers for the downlink (2) and uplink (4) problems. However, the final step of this algorithm in (10) requires a global knowledge about the CSI which makes its distributed implementation difficult, especially when dealing with a large number of users and antennas.

IV. Decentralized approach for large dimension system

In this section we introduce our decentralized algorithm for a system with large dimensions based on the approximated ICI thresholds. It is known that the growing dimensions of a random matrix results in some deterministic behaviors about the distribution of its eigenvalues that can be utilized for processing simplifications [15]. Authors in [9] have considered a system with large dimensions, i.e., large number of users and antennas. In addition, they assume that the variances of channel entries are scaled by the number of antennas. This assumption allows deriving a large dimension approximation for (5), (9) and (10) that results in a decentralized beamforming approach which relies on local CSI and the average statistics of the other channels. According to [9], an approximation for the optimal uplink power defined by (5) can be formulated as

$$\lambda_k = ((1+\frac{1}{\gamma_k})(\frac{a^2_{b_k,k} m_{\boldsymbol{\Sigma}_{b_k}}(-1)}{1+a^2_{b_k,k}\lambda_k m_{\boldsymbol{\Sigma}_{b_k}}(-1)}))^{-1} \tag{11}$$

where, $a^2_{b_k,k}$ is the pathloss from the k^{th} user to its serving base station. $\boldsymbol{\Sigma}_{b_k}$ is defined by (6) and $m_{\boldsymbol{\Sigma}_{b_k}}(-1)$ is the Stieltjes transform of the Gram matrix $\boldsymbol{\Sigma}_{b_k}$ at point $z=-1$ as defined by Theorem A-2 in the appendix.

Similarly, the matrix inversion for δ_k values can be approximated by [9]

$$G_{i,j} = \begin{cases} \frac{1}{\gamma_i}(\frac{a^2_{b_i,i} m_{\boldsymbol{\Sigma}_{b_i}}(-1)}{\eta_{b_i,i}})^2 & i=j \\ \frac{-1}{N_{\text{a}}}\frac{a^2_{b_j,i}a^2_{b_j,j} m'_{\boldsymbol{\Sigma}_{b_j}}(-1)}{\eta^2_{b_j,i}\eta^2_{b_j,j}} & i\neq j \end{cases} \tag{12}$$

where

$$\eta_{b_j,i} = 1 + a^2_{b_j,i}\lambda_i m_{\boldsymbol{\Sigma}_{b_j}}(-1) \tag{13}$$

and where $m'_{\boldsymbol{\Sigma}_b}(-1)$ is the differential of $m_{\boldsymbol{\Sigma}_b}(z)$ with respect to z at point $z=-1$ (see the Appendix).

Under the large dimension assumption, the beamforming vectors in (8) can be found using locally acquired channel knowledge and approximated dual uplink powers (11) and cross-coupling matrix (12) instead of (5) and (9), resulting in a distributed beamforming algorithm. However, the problem with the approximated method is that it cannot guarantee the target SINRs for finite number of antennas as the error in approximations is translated into variations in the resulted SINRs which can be less or more than the target SINRs.

A. Approximation of intercell interference terms

The method proposed here relies on approximately optimal ICI values instead of approximated uplink powers as in [9]. The approximate ICI values remain valid for a given set of users until a change occurs in the statistics of the channel, i.e., when a user changes its location. This leads potentially to a significant reduction of the required backhaul signaling depends only on the large scale parameters, i.e. pathloss between each BS and active node.

The large dimension approximation for ICI terms can be achieved by using (11), (12) and (8). From (2) and (3), it is clear that the intercell interference from all the base stations towards user k is

$$\sum_{b\neq b_k}\epsilon^2_{b,k} = \sum_{b\in\overline{\mathcal{B}_k}}\sum_{l\in\mathcal{U}_b}|\mathbf{h}_{b,k}\mathbf{w}_{b,l}|^2 \tag{14}$$

where, the intercell interference term from the bth base station to user k is,

$$\epsilon^2_{b,k} = \sum_{l\in\mathcal{U}_b}|\mathbf{h}_{b,k}\mathbf{w}_{b,l}|^2 \tag{15}$$

Considering (8), the intercell interference term in (15) can be written as follows,

$$\epsilon^2_{b,k} = \sum_{l\in\mathcal{U}_b}\sqrt{\delta_{b,l}}|\mathbf{h}_{b,k}\hat{\mathbf{w}}_{b,l}|^2 \tag{16}$$

where $\delta_{b,l}$ values can be found from (10) and the approximation for the cross-terms $|\mathbf{h}_{b,i}\hat{\mathbf{w}}_{b,j}|^2$ are defined by (12),

$$|\mathbf{h}_{b,i}\hat{\mathbf{w}}_{b,j}|^2 \approx \frac{1}{N_{\text{a}}}\frac{a^2_{b_j,i}a^2_{b_j,j} m'_{\boldsymbol{\Sigma}_{b_j}}(-1)}{\eta^2_{b_j,i}\eta^2_{b_j,j}} \quad i\neq j \tag{17}$$

Therefore, the ICI from the b^{th} BS to the k^{th} user can be written as,

$$\epsilon^2_{b,k} = \sum_{l\in\mathcal{U}_b}\sqrt{\delta_{b,l}}\frac{1}{N_{\text{a}}}\frac{a^2_{b_l,k}a^2_{b_l,l} m'_{\boldsymbol{\Sigma}_{b_l}}(-1)}{\eta^2_{b_l,k}\eta^2_{b_l,l}} \tag{18}$$

This approximation allows derivation of approximate optimal ICI based on statistics of the user channels. Each BS needs knowledge about user specific average statistics, i.e., pathloss values from other BSs based on which each BS can locally and independently calculate the approximately optimal ICI values.

B. Distributed beamforming based on approximated ICI values

Using any fixed ICI value in (2) is a special case that results in a suboptimal performance in general. In [8], [10], [11], an agreement on optimal fixed ICI values is achieved via exchange of scalar ICI parameters, i.e., local copies of ICI terms or corresponding dual variables. Another straightforward decentralized approach is to enforce all inter-cell interference to zero [8]. In all cases, however, the intra-cell interference between local users can be optimally handled. Solving (2) with the approximated ICI values $\epsilon_{b,k}$ developed in the previous subsection leads to an algorithm that benefits from both of the locally optimal beamforming design and near optimal ICI knowledge. This property brings significant gains compared to other suboptimal methods like inter-cell interference nulling. The proposed algorithm is summarized in Algorithm 1.

Algorithm 1 Approximation of the ICI values.

```
Initialize the ICI values based on the exchanged pathloss values
loop
  if Any change in the user statistics then
    Exchange the updated pathloss values a_{b,k} among coupled BSs.
    Update the approximated λ_k values, m_{Σ_{b_k}}(−1) and its derivative from (11) and Theorem A-2.
    Get approximated δ values from (10).
    Update the approximated ICIs based on (18).
  end if
  Use the approximated ICIs as a fixed ε²_{b,k} in (2) and solve the subproblems locally for getting the optimal downlink beamformers.
end loop
```

The local problems can be solved by reformulating (2) as BS specific SOCP or solved iteratively as in [11]. The proposed algorithm guarantees the target SINRs because the feasible solution of the optimization problem defined by (2) always satisfies the constraints and the possible error in approximations is translated into a somewhat higher transmit power at BSs compared to the optimal centralized solution.

V. Numerical Analysis

The algorithm developed in the previous section satisfies the target SINRs for all users; however, the error in approximations results a higher transmit power at BSs. In order to evaluate the difference between the optimal transmit power and the power resulted from the approximated algorithm, an extensive multi-cell simulation study is carried out in this section. A network with 7 cells is considered and users are scattered on the coverage area of the network, in a way that each cell contains 4 users. Exponential pathloss model is used for assigning the pathloss to each user.

$$a_{b,k} = (\frac{d_0}{d_{b,k}})^2 \tag{19}$$

where $d_{b,k}$ is distance between base station b and user k. The pathloss exponent is 2 and the reference distance (d_0) is 1m. The pathloss from a base station to the boundary of the reference distance of the neighboring base station is 50dB. The users are dropped randomly for each trial and in total 1000 user drops are used for calculating the average transmit power. Fig. 1 depicts the network with 7 cells and the users dropped on the coverage area.

Figs. 2 and 3 illustrate the transmit powers versus the number of antennas for 0dB and 10dB SINR target, respectively. The fading characteristics per antenna is i.i.d. It is clear that the gap between the approximate and optimal algorithm (denoted as SOCP) diminishes as the number of antennas increases. When the number of antennas is equal to 28 the gap is less

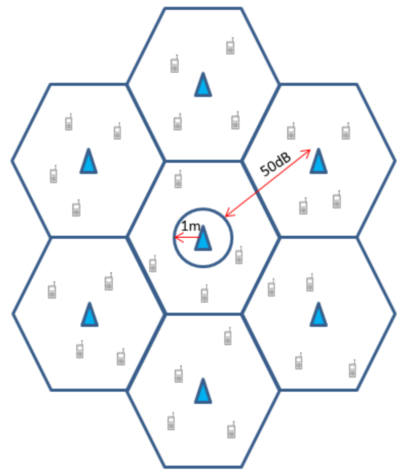

Fig. 1: A Network with 7 Cells

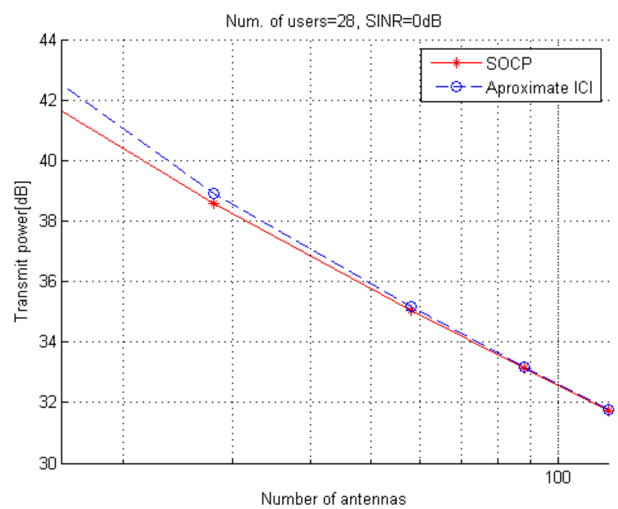

Fig. 2: Required transmit power for 0 dB SINR target.

than 0.5dB which indicates that the approximate algorithm provides a good solution for the practical scenarios with a limited number of antennas. The gap for the case with 0dB SINR in Fig. 3 does not exceed 1dB even when number of antennas is smaller than 28, however, for the case with 10dB SINR target, the approximated case becomes infeasible.

In Fig. 4, the transmit powers resulting from other sub-optimal methods, such as matched filter (MF) and zero-forcing (ZF) are compared with the optimal centralized case and the approximately optimal case. From this figure it is clear that SOCP algorithm and the approximate ICI algorithm outperform the ZF over a wide range of number of antennas even in the ideal case with i.i.d fading statistics. The gap in performance is mainly due to the fact that the ZF algorithm wastes a degree of freedom for nulling the interference towards the distant users while the SOCP algorithm finds the optimal balance between interference suppression and maximizing the desired signal level. MF beamforming must be dealt with more care since the SINR target can be guaranteed only asymptotically, i.e., when the number of antennas approaches infinity. The increasing curve in Fig. 4 shows the resulting

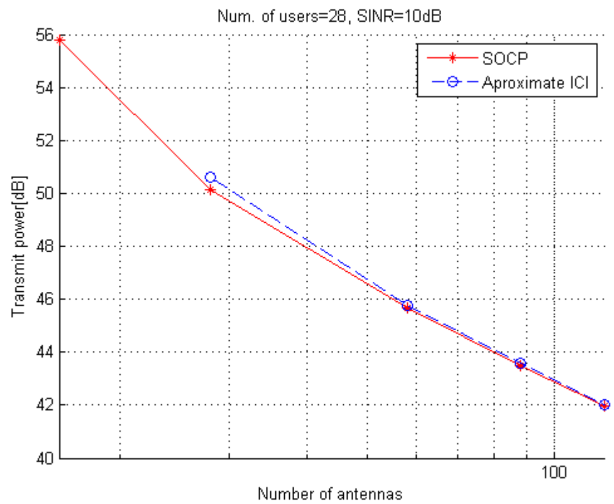

Fig. 3: Required transmit power for 10 dB SINR target

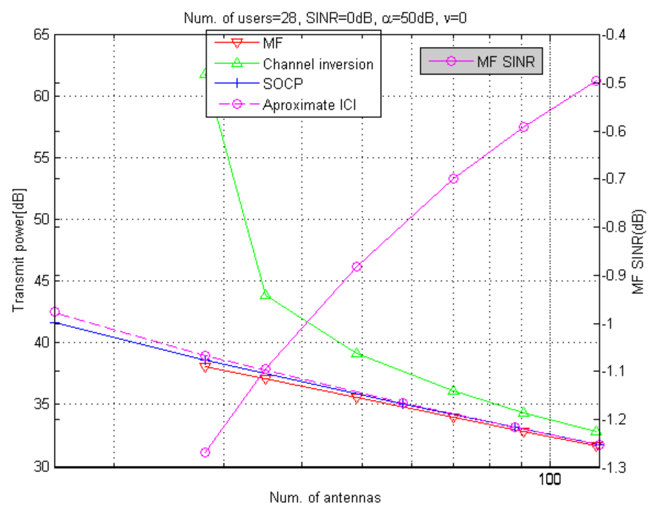

Fig. 4: Comparison of required transmit power for 0 dB SINR target

SINR when simple MF beamforming is used. For a small number of antennas, the resulted SINR is well below the target and by increasing the number of antennas it approaches the target SINR. Nevertheless, the achieved SINR is 0.5 dB below the target SINR even for 120 antennas at each base station even though the transmit powers for the optimal and the MF beamforming are almost the same. The gap between ZF and SOCP schemes also depends on target SINR. For high target SINR values, the optimal strategy approaches the ZF beamforming. The performance difference decreases fast in i.i.d channel when the number of antennas increases as shown in Fig. 5.

Figs. 2–5 consider an idealistic scenario with i.i.d. fading characteristics per antenna. In practice, however, non-zero correlation and coupling among antenna elements, imperfect CSI and hardware, non-ideal medium, etc. must be taken into account in order to make a realistic comparison. In

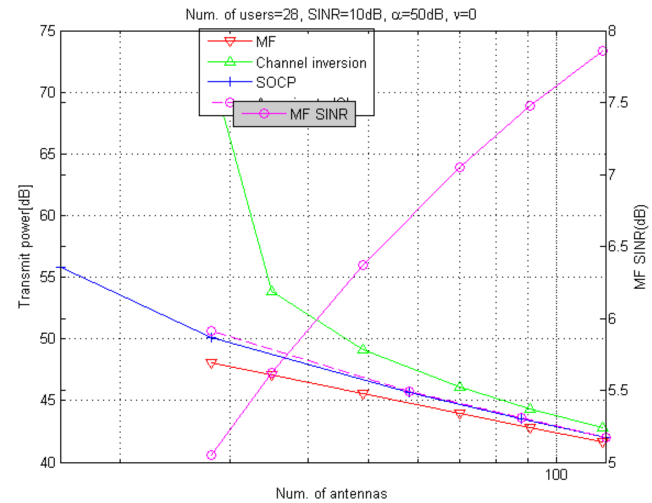

Fig. 5: Comparison of required transmit power for 10 dB SINR target

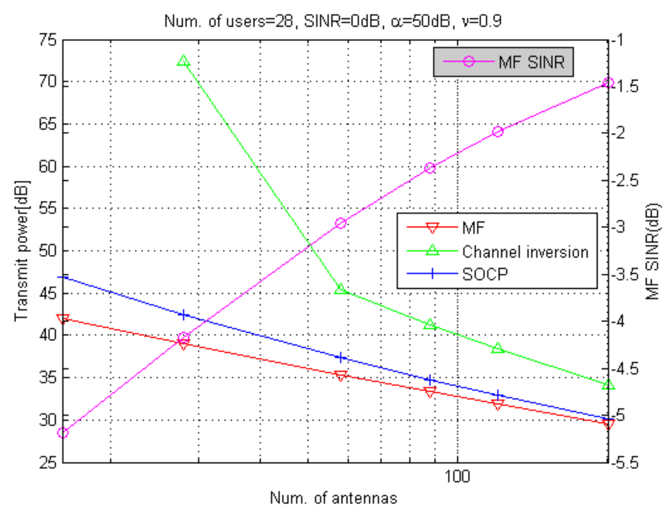

Fig. 6: Comparison of required transmit power for 0 dB SINR target with $\rho = 0.9$ correlation between adjacent antennas.

Fig. 6, the effect of correlation among antenna elements is considered by using a simple Kronecker channel model with correlation factor 0.9 between adjacent antenna elements. It can be seen from Fig. 6, the correlation increases the gap between SOCP and ZF approaches even further. Therefore, the use of advanced beamforming schemes can be justified even when the size of antenna array becomes large.

VI. Conclusions

Intercell interference is a key parameter in the design of distributed beamforming algorithm as it couples the subproblems at base stations. In this work, a large dimension approximation for the optimal ICI has been considered. According to this approximation an algorithm has been proposed for decoupling the subproblems at base stations which results in a significant reduction in backhaul information exchange rate and processing load. This algorithm guarantees the target SINRs without any major loss of performance as compared to the optimal centralized design as the dimensions of the system grow large. Massive MIMO with a practically limited array size is an example of a system where the proposed algorithm can be utilized.

APPENDIX

For the convenience of the reader, some important lemmas and theories from random matrix theory are briefly summarized. At first Stieltjes transform is introduced:

Definition A-1. [15], Assuming a real valued bounded measurable function over $\mathbb{R}$ denoted by F, the Stieltjes transform of F for $z \in Supp(F)^c$ is defined as:

$$m_F(z) = \int_{-\infty}^{\infty} \frac{1}{\lambda - z} dF(\lambda) \tag{20}$$

F can be any function that satisfies the specified condition in Definition A-1. However, for studies here, F is the empirical spectral distribution (e.s.d.) of a $N \times N$ Hermitian matrix $\mathbf{Y}_N$ defined for $x \in \mathbb{R}$ as,

$$F(x) = \frac{1}{N} \sum_{j=1}^{N} 1_{\lambda_j \leq x}(x) \tag{21}$$

Where $\lambda_1, ..., \lambda_N$ are eigenvalues of $\mathbf{Y}_N$. $1_{\lambda_j \leq x}$ is the indicator function which gives 1 when λ_j is less than x and 0 otherwise.

Then, for the hermitian matrix $\mathbf{Y}_N$ it is known that,

$$m_F(z) = \frac{1}{N} tr(\mathbf{Y} - z\mathbf{I}_{\mathrm{N}})^{-1} \tag{22}$$

The following theorem gives the equations for deriving an equivalent of Stieltjes transform of Gram matrix $\mathbf{Y}\mathbf{Y}^{\mathrm{H}}$.

Theorem A-2. [9] [16], Consider a $N \times n$ random matrix denoted by $\mathbf{Y}$, such that elements of $\mathbf{Y}$ are independent and zero mean. The variance of entries is given by $E[|y_{i,j}|^2] = \alpha_{i,j}^2$ which are uniformly bounded from above. Also some soft restrictions are assumed over higher moments. then,

$$\frac{1}{N} tr(\mathbf{Y}\mathbf{Y}^{\mathrm{H}} - z\mathbf{I}_{\mathrm{N}})^{-1} - \frac{1}{N} tr(\mathbf{\Theta}(z)) \xrightarrow{n \to \infty, \frac{N}{n} \to c} 0, \forall z \in \mathbb{C} - \mathbb{R}^+ \tag{23}$$

Where $\mathbf{\Theta}(z) = diag(\theta_1(z), ..., \theta_N(z))$ is a deterministic matrix valued function, analytic in $\mathbb{C} - \mathbb{R}^+$. The entries of this function can be found by initializing and iterating the following system of $N + n$ equations:

$$\theta_i(z) = \frac{-1}{z(1 + \frac{1}{n} \sum_{j=1}^{n} \alpha_{i,j}^2 \hat{\theta}_j(z))} \; \forall 1 \leq i \leq N \tag{24}$$

$$\hat{\theta}_j(z) = \frac{-1}{z(1 + \frac{1}{n} \sum_{i=1}^{N} \alpha_{i,j}^2 \theta_i(z))} \; \forall 1 \leq j \leq n \tag{25}$$

The derivative of Stieltjes transform of the matrix $\mathbf{Y}\mathbf{Y}^{\mathrm{H}}$ is given by $\frac{1}{N} tr(\mathbf{\Theta}'(z))$. Where $\mathbf{\Theta}'(z) = diag(\theta'_1(z), ..., \theta'_N(z))$. The entries can be found by solving a system of equation resulted from taking derivatives of (24) and (25) with respect to z [9].

REFERENCES

[1] M. Karakayali, G. J. Foschini, R. A. Valenzuela, and R. D. Yates, "On the maximum common rate achievable in a coordinated network," in *IEEE International Conference on Communications, ICC'06.*, vol. 9. IEEE, 2006, pp. 4333–4338.

[2] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network mimo: Overcoming intercell interference in indoor wireless systems," in *Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*. IEEE, 2007, pp. 83–87.

[3] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, 2011.

[4] Y. Gao, Y. Li, H. Yu, and S. Gao, "Performance of dynamic comp cell selection in 3gpp lte system level simulation," in *IEEE 3rd International Conference on Communication Software and Networks (ICCSN)*. IEEE, 2011, pp. 210–213.

[5] L. Venturino, N. Prasad, and X. Wang, "Coordinated scheduling and power allocation in downlink multicell ofdma networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 6, pp. 2835–2848, 2009.

[6] P. Komulainen, A. Tölli, and M. Juntti, "Effective CSI signaling and decentralized beam coordination in TDD multi-cell MIMO systems," *IEEE Trans. Signal Processing*, vol. 61, no. 9, pp. 2204–2218, May 2013.

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

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

[9] S. Lakshminaryana, J. Hoydis, M. Debbah, and M. Assaad, "Asymptotic analysis of distributed multi-cell beamforming," in *21st IEEE nternational Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2010, pp. 2105–2110.

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

[11] ——, "Multi-cell beamforming with decentralized coordination in cognitive and cellular networks," *IEEE Trans. Signal Processing (to appear)*, 2013.

[12] 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.

[13] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Mag.*, vol. 30, no. 1, pp. 40 – 60, Jan. 2013.

[14] C. Shen, T.-H. Chang, K.-Y. Wang, Z. Qiu, and C.-Y. Chi, "Distributed robust multicell coordinated beamforming with imperfect csi: An admm approach," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2988–3003, 2012.

[15] M. D. Romain Couillet, *Random Matrix Methods for Wireless Communications*, 1st ed., 2011.

[16] P. L. Walid Hachem and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *Ann. Appl. Probab. Volume 17, Number 3*, 2007.