Distributed Precoder Design Under Per-Small Cell Power Constraint

Maik Röper†, Patrick Svedman‡ and Armin Dekorsy†
† Department of Communications Engineering, University of Bremen, Bremen, Germany
‡ ZTE Wistron Telecom AB, Stockholm, Sweden
Email: {roeper,dekorsy}@ant.uni-bremen.de, patricksvedman@zte.com.cn

***Abstract*—In this paper, a novel distributed precoding (DiP) algorithm for ultra-dense small cell (SC) networks is developed, where the SCs cooperate to perform a joint transmission to users (UEs) with limited and individual transmit powers. Different to most state of the art (SotA) DiP algorithms, the proposed precoder design is based on the assumption that each SC has only local channel state information (CSI) available. Additionally, there are no constraints on the number of antennas for each SC, but only one constraint on the sum of all transmit antennas. A solution for the considered problem based on the Lagrangian method of multipliers (MoM) is presented, by formulating the precoder design as a constrained convex optimization problem. The obtained solution can be implemented in a fully distributed way among the SCs by using the preconditioned Richardson (PR) iteration. In numerical simulations, the convergence of the proposed DiP algorithm is verified and it is shown that the sum rate significantly increases, if the SCs cooperate with each other.**

I. Introduction

The deployment of dense small cell (SC) networks in new generations of wireless communication systems has attracted growing interests in recent years [1], [2]. As an advancement evolving in dense SC networks, multi-input multi-output (MIMO) technology can significantly improve system performance in coverage, capacity and energy efficiency [3]. Whereas, the inter cell interference (ICI) caused by the multicell MIMO channels needs to be coped with. Therefore, the downlink (DL) transmission cooperation among SCs is focused on in this paper, which is usually executed in a central unit with user data and channel state information (CSI) sharing over a backhaul (BH) network [4]. However, considering limited BH capacity and high complexity for the centralized precoding, the distributed cooperative precoding is in particular of interest, which can provide considerable data rate improvement by using interference coordination strategies such as coordinated power control, interference alignment or coordinated joint transmission [5], [6].

In [7] a distributed implementation of the centralized zero forcing (ZF) and minimum mean square error (MMSE) precoder under a sum power constraint (SPC) [8] is shown, where an In-Network processing (INP) technique [9] is applied for the cooperative processing among the distributed SCs. However, the SPC for the precoder design is usually not of practical interest. Instead, the per-SC power constraint or per-antenna power constraint is normally used. E.g., in [10], [11], [12], several distributed precoding schemes are proposed to minimize the ICI achieving maximization of signal to interference plus noise ratio (SINR) or system sum rate. Moreover, MSE based distributed precoders under per-base station (BS) power constraint have also been developed in [13], [14], [15], where the local precoding and receive filtering matrices are updated in an alternated fashion. Nevertheless, the calculation of the optimal receive filtering matrices at each SC leads to high latency due to the computational complexity and therefore, suboptimal precoders are more favorable in practice.

In this paper, we present a new suboptimal distributed linear precoder under per-SC power constraint (PSPC) for joint transmission by solving a constrained convex optimization problem where each SC has only local CSI available. Compared to the sequential update of the distributed precoder [15], the local precoders are updated in a parallel way here, and information is only required to be exchanged between SCs during an iterative process. Similar to [7], the development of the distributed precoding algorithm is also inspired by INP techniques. One numerical method, the preconditioned Richardson (PR) iteration [16], is applied for a distributed implementation of the precoding algorithm among the SCs. Moreover, following the same principle of the DiP algorithm in [7], the distributed precoding approach can also be specified into the update of the local precoding matrix or the update of the local transmit signal, depending on the system scenario. Here, we only focus on the distributed update of local precoding matrix under PSPC. More discussion and evaluation of the proposed algorithm will be detailed in the following sections.

The remainder of this paper is structured as follows: The system model and objective problem are described in Section II. Then, we introduce and discuss the approach for developing the DiP algorithm under PSPC in Section III. In the subsequent Section IV, the performance of the proposed algorithm is investigated and evaluated through numerical simulations. Finally, the paper is concluded in Section V.

II. Preliminaries

A. System Model

We consider joint transmission in a small cell network where N_{SC} SCs are connected via perfect inter-SC links and cooperatively serve N_{UE} UEs. Although ideal links between the SCs are assumed, limitations in terms of inter-SC links

978-1-5386-6358-5/18/$31.00 ©2018 IEEE

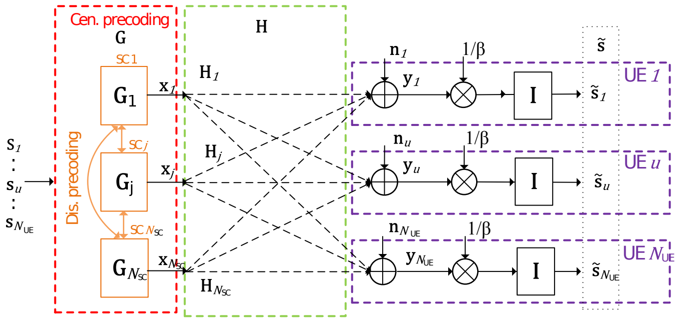

Fig. 1. Diagram of signal processing in central/distributed precoding

capacity and latency are considered in section IV by assuming partially coordinated transmission and limiting the number of iterations. Each UE and SC is equipped with N_R receive and N_T transmit antennas, respectively, leading to a general $N_\mathrm{I} \times N_\mathrm{O}$ MIMO system with $N_\mathrm{I} = N_\mathrm{SC}N_\mathrm{T}$ transmit and $N_\mathrm{O} = N_\mathrm{UE}N_\mathrm{R}$ receive antennas in total and $N_\mathrm{I} > N_\mathrm{O}$. Even if each SC sees an undetermined $N_\mathrm{T} \times N_\mathrm{O}$ system, a joint solution for the overdetermined $N_\mathrm{I} \times N_\mathrm{O}$ MIMO system can be found without the need of any centralized processing unit by using cooperation between the SCs.

As shown in Fig. 1, the system input vector $\mathbf{s}$ of length N_O is composed of all UE data vectors with alphabet elements, i.e., $\mathbf{s} = [\mathbf{s}_1^\mathrm{T}, .., \mathbf{s}_{N_\mathrm{UE}}^\mathrm{T}]^\mathrm{T} \in \mathbb{A}^{N_\mathrm{O}\times 1}$ with $\mathrm{E}\left\{\mathbf{s}\mathbf{s}^H\right\} = \mathbf{I}_{N_\mathrm{O}}$, where $\mathbf{I}_{N_\mathrm{O}}$ denotes the identity matrix of dimension N_O. Besides, we assume for now that the complete data vector $\mathbf{s}$ is available at each SC, e.g., by downloading or caching from the cloud [17]. For each time instance, the vector $\mathbf{s}$ is precoded by the local precoding matrices $\mathbf{G}_j \in \mathbb{C}^{N_\mathrm{T}\times N_\mathrm{O}}$ of SC $j \in \{1, .., N_\mathrm{SC}\}$, and the precoded signal vectors $\mathbf{x}_j = \mathbf{G}_j\mathbf{s} \in \mathbb{C}^{N_\mathrm{I}\times 1}$ are then transmitted to the UEs simultaneously. At the receiver side, each UE u receives a signal $\mathbf{y}_u \in \mathbb{C}^{N_\mathrm{O}\times 1}$, which is the superposition from all N_SC transmitted signals $\mathbf{x}_j$ with complex additive white Gaussian noise $\mathbf{n}_u \sim \mathcal{CN}(0, \sigma_n^2\mathbf{I}_{N_\mathrm{R}})$, assuming that the noise power σ_n^2 is identical for all UEs. Then, each UE re-scales the received signals with a factor β which is determined at the transmitter considering the power constraint and can be easily obtained or estimated at the receiver side.[1] Thus, the recovered signal $\tilde{\mathbf{s}}_u$ at UE u is given by

$$\tilde{\mathbf{s}}_u = \frac{1}{\beta}\mathbf{y}_u = \frac{1}{\beta}\sum_{j=1}^{N_\mathrm{SC}} \mathbf{H}_{uj}\mathbf{G}_j\mathbf{s} + \mathbf{n}_u, \tag{1}$$

where $\mathbf{H}_{uj}$ denotes the channel between UE u and SC j, which is assumed to be locally known by SC j.

B. Problem Formulation

Let vector $\tilde{\mathbf{s}}$ be the total system output, i.e., $\tilde{\mathbf{s}} = [\tilde{\mathbf{s}}_1^\mathrm{T}, .., \tilde{\mathbf{s}}_{N_\mathrm{UE}}^\mathrm{T}]^\mathrm{T}$, then we aim to minimize the mean square error between the total recovered signals $\tilde{\mathbf{s}}$ and the original UE data $\mathbf{s}$ by a proper design of precoding matrices $\mathbf{G}_j$ as well as the scaling factor β under a limited transmit power for each SC j. The corresponding optimization problem of our system reads as

$$\begin{aligned} \min_{\mathbf{G}_j,\beta}\ & \mathrm{E}\left\{\|\mathbf{s} - \tilde{\mathbf{s}}\|^2\right\} \\ \text{s.t.}\ & \mathrm{E}\left\{\|\mathbf{G}_j\|_F^2\right\} \leq P_j,\ \forall\, j = 1, .., N_\mathrm{SC} \end{aligned} \tag{2}$$

where P_j is the maximum available transmit power of SC j. In the following, we will introduce our novel approach for developing the local precoder under PSPC in a distributed way.

III. Distributed Precoder Design

In this section, we adopt the Lagrangian method of multipliers (MoM) [18] to solve the constrained optimization problem (2) and apply the PR iteration to distribute the processing among the SCs achieving the PR based distributed precoding (PR-DiP) algorithm. We first rewrite the total recovered signals $\tilde{\mathbf{s}}$ in a matrix vector product form:

$$\tilde{\mathbf{s}} = \frac{1}{\beta}\left(\mathbf{H}\mathbf{G}\mathbf{s} + \mathbf{n}\right), \tag{3}$$

where the UE data vector $\mathbf{s}$ is precoded by the stacked precoding matrix $\mathbf{G} = [\mathbf{G}_1^\mathrm{T}, .., \mathbf{G}_j^\mathrm{T}, ..., \mathbf{G}_{N_\mathrm{SC}}^\mathrm{T}]^\mathrm{T}$ and the entire channel matrix $\mathbf{H} = [\mathbf{H}_1, ..., \mathbf{H}_{N_\mathrm{SC}}]$ consists of the local channel matrices $\mathbf{H}_j = [\mathbf{H}_{1j}^\mathrm{T}, ..., \mathbf{H}_{N_\mathrm{UE}j}^\mathrm{T}]^\mathrm{T}$ of each SC j. At each SC $j = 1, ..., N_\mathrm{SC}$, the local channel matrix $\mathbf{H}_j$ is assumed to be perfectly known and to stay constant during the whole processing time for designing the distributed precoder. If only erroneous CSI is available, but the statistics of the CSI error is known, a bayesian approach [19], [20] can be applied to increase the robustness of the PR-DiP algorithm against erroneous CSI. The design of a robust distributed precoder is not considered in this paper and left for future study.
The total noise vector $\mathbf{n}$ is stacked as $\mathbf{n} = [\mathbf{n}_1^\mathrm{T}, , ..., \mathbf{n}_{N_\mathrm{UE}}^\mathrm{T}]^\mathrm{T}$. The system objective problem (2) then becomes:

$$\begin{aligned} \min_{\mathbf{G},\beta}\ & \mathrm{E}\left\{\left\|\mathbf{s} - \beta^{-1}\Big(\mathbf{H}\mathbf{G}\mathbf{s} + \mathbf{n}\Big)\right\|^2\right\} \\ \text{s.t.}\ & \mathrm{E}\left\{\|\mathbf{G}_j\|_F^2\right\} \leq P_j,\ \forall\ j = 1, .., N_\mathrm{SC}. \end{aligned} \tag{4}$$

By applying the MoM we get the Lagrangian function $\mathcal{L}(\cdot)$ for (4) as

$$\begin{aligned} \mathcal{L}(\mathbf{G}, \beta, \mathbf{\Lambda}) =\ & \mathrm{tr}\left\{\left(\beta^{-1}\mathbf{H}\mathbf{G}\mathbf{s} - \mathbf{s}\right)\left(\beta^{-1}\mathbf{H}\mathbf{G}\mathbf{s} - \mathbf{s}\right)^\mathrm{H} + \beta^{-2}\sigma_n^2\mathbf{I}_{N_\mathrm{O}}\right\} \\ & + \mathrm{tr}\left\{\mathbf{G}\mathbf{\Lambda}\mathbf{G}^\mathrm{H}\right\} - \mathbf{p}^\mathrm{H}\mathbf{\Lambda}\mathbf{p}, \end{aligned} \tag{5}$$

where $\mathrm{tr}\{\cdot\}$ denotes the trace operator and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, .., \lambda_{N_\mathrm{SC}}) \otimes \mathbf{I}_{N_\mathrm{T}}$ is a diagonal matrix consisting of the Lagrangian multipliers λ_j corresponding to the power constraints in (2). In addition, we also define a stacked vector $\mathbf{p} = [\sqrt{P_1}, .., \sqrt{P_{N_\mathrm{SC}}}]^\mathrm{T}$ that consists of the power constraint scalars P_j for each SC j.

Moreover, for the constrained optimization problem (4), according to the Karush-Kuhn-Tucker (KKT) conditions [21], we have:

$$\frac{\partial}{\partial\mathbf{G}}\mathcal{L}(\mathbf{G}^*, \beta^*, \mathbf{\Lambda}^*) = \mathbf{0} \tag{6}$$

[1]Since we focus on the development of a precoder, the post processing at UEs is simplified (denoted by a block with the identity matrix $\mathbf{I}$ in Fig. 1). Nevertheless, independent filtering at each UE could also be implemented to further increase the performance, which is not considered here.

$$\frac{\partial}{\partial \beta}\mathcal{L}(\mathbf{G}^*, \beta^*, \mathbf{\Lambda}^*) = 0 \tag{7}$$

$$\mathrm{tr}\left\{\mathbf{G}_j^*(\mathbf{G}_j^*)^{\mathrm{H}}\right\} - P_j \leq 0,\ j = 1, .., N_{\mathrm{SC}} \tag{8}$$

$$\lambda_j^*\left(\mathrm{tr}\left\{\mathbf{G}_j^*(\mathbf{G}_j^*)^{\mathrm{H}}\right\} - P_j\right) = 0, \lambda_j^* \geq 0. \tag{9}$$

where $\mathbf{G}^*, \beta^*, \mathbf{\Lambda}^*$ are the optimal solutions for the considered optimization problem. To the best of our knowledge, no analytic closed form solution for the equation system (6)-(9) can be given. Hence, it has to be solved in an iterative fashion. If $\beta^{(k)}$ and $\mathbf{\Lambda}^{(k)}$ are kept fixed at iteration k, the updated central precoding matrix is achieved by solving:

$$\frac{\partial}{\partial \mathbf{G}}\mathcal{L}(\mathbf{G}^{(k+1)}, \beta^{(k)}, \mathbf{\Lambda}^{(k)}) = \mathbf{0} \Leftrightarrow \tag{10}$$

$$\mathbf{G}^{(k+1)} = \left(\mathbf{H}^{\mathrm{H}}\mathbf{H} + {\beta^{(k)}}^2\mathbf{\Lambda}^{(k)}\right)^{-1}\beta^{(k)}\mathbf{H}^{\mathrm{H}} \tag{11}$$

However, the update of $\mathbf{G}^{(k+1)}$ cannot be decomposed among SCs due to the matrix inverse. Therefore, we apply the PR method to achieve another iterative update of $\mathbf{G}$ from (11), which can be distributed among SCs in parallel. The update of $\mathbf{G}^{(k+1)}$ is then given by

$$\mathbf{G}^{(k+1)} = \mathbf{G}^{(k)} - \omega\mathbf{T}\Bigg(\left(\mathbf{H}^{\mathrm{H}}\mathbf{H} + {\beta^{(k)}}^2\mathbf{\Lambda}^{(k)}\right)\mathbf{G}^{(k)} - \beta^{(k)}\mathbf{H}^{\mathrm{H}}\Bigg), \tag{12}$$

where $\omega > 0$ is the stepsize that needs to be properly chosen to ensure the convergence [16]. The preconditioning matrix $\mathbf{T}$ is chosen to be

$$\mathbf{T} = \begin{bmatrix} \left(\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1 + {\beta^{(k)}}^2\lambda_1^{(k)}\mathbf{I}\right)^{-1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \left(\mathbf{H}_{N_{\mathrm{SC}}}^{\mathrm{H}}\mathbf{H}_{N_{\mathrm{SC}}} + {\beta^{(k)}}^2\lambda_{N_{\mathrm{SC}}}^{(k)}\mathbf{I}\right)^{-1} \end{bmatrix}. \tag{13}$$

Due to the block diagonal structure of $\mathbf{T}$ and the linearity of the system (12), we can decompose the centralized update of $\mathbf{G}^{(k+1)}$ into parallel updates of local precoding matrices $\mathbf{G}_j^{(k+1)}$ among SCs $j = 1, .., N_{\mathrm{SC}}$:

$$\begin{aligned}\mathbf{G}_j^{(k+1)} =\ & \mathbf{G}_j^{(k)} - \omega\left(\mathbf{H}_j^{\mathrm{H}}\mathbf{H}_j + {\beta^{(k)}}^2\lambda_j^{(k)}\mathbf{I}_{N_{\mathrm{T}}}\right)^{-1} \\ & \cdot\left(\mathbf{H}_j^{\mathrm{H}}\sum_{i=1}^{N_{\mathrm{SC}}}\mathbf{H}_i\mathbf{G}_i^{(k)} + {\beta^{(k)}}^2\lambda_j^{(k)}\mathbf{G}_j^{(k)} - \beta^{(k)}\mathbf{H}_j^{\mathrm{H}}\right)\end{aligned} \tag{14}$$

For each SC j, the precoding matrix $\mathbf{G}_j^{(k+1)}$ is updated locally, whereas in each iteration some additional informations $\mathbf{H}_i\mathbf{G}_i^{(k)}$ from all other SCs $i \neq j$ are required, which needs to be delivered over inter-SC links.

After the exchange of the local information $\mathbf{H}_j\mathbf{G}_j^{(k)}$ between the SCs, the optimal scaling factor $\beta^{(k)}$ for fixed $\mathbf{G}^{(k)}$ can be calculated by solving $\frac{\partial}{\partial\beta}\mathcal{L}(\mathbf{G}^{(k)}, \beta^{(k)}, \mathbf{\Lambda}^{(k)}) = 0$ at each SC independently. Since all the information is available at each SC j, the scaling factor $\beta^{(k)}$ at iteration k is given by

$$\beta^{(k)} = \frac{\mathrm{tr}\left\{\mathbf{H}\mathbf{G}^{(k)}(\mathbf{G}^{(k)})^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\right\} + N_{\mathrm{O}}\sigma_n^2}{\mathrm{Re}\left\{\mathrm{tr}\left\{\mathbf{H}\mathbf{G}^{(k)}\right\}\right\}}$$

$$= \frac{\left\|\sum_{j=1}^{N_{\mathrm{SC}}}\mathbf{H}_j\mathbf{G}_j^{(k)}\right\|_F^2 + N_{\mathrm{O}}\sigma_n^2}{\mathrm{Re}\left\{\mathrm{tr}\left\{\sum_{j=1}^{N_{\mathrm{SC}}}\mathbf{H}_j\mathbf{G}_j^{(k)}\right\}\right\}}. \tag{15}$$

Note that the scaling factor $\beta^{(k)}$ can be updated locally and is the same over all SCs for each iteration k.

An analytic expression for optimal Lagrangian multipliers λ_j cannot be given. However, the update for the local precoding matrices $\mathbf{G}_j^{(k+1)}$ in (14) is depending on $\lambda_j^{(k)}$, so the power constraints after the upcoming iteration $k+1$ can be ensured by numerically solving

$$\left\|\mathbf{G}_j^{(k+1)}\right\|_F^2 = P_j\,. \tag{16}$$

At each SC j the update of its precoding matrix $\mathbf{G}_j^{(k+1)}$ only depends on the single multiplier $\lambda_j^{(k)}$ and therefore, it can be found by using a root-finding algorithm, e.g., Newton's or bisection method, without further information exchange. Due to the KKT condition (9), the values of multipliers λ_j have to be non-negative. If the solutions for (16) are negative, the multipliers should be set to zero.

In Algorithm 1, our PR-DiP algorithm under PSPC is summarized.

Algorithm 1 PR-DiP under PSPC

For each SC $j = 1, ..., N_{\mathrm{SC}}$ in parallel

1: **Initialization**: local precoding matrix $\mathbf{G}_j^{(0)} = \nu_j(\mathbf{H}_j^{\mathrm{H}}\mathbf{H}_j)^{-1}\mathbf{H}_j^{\mathrm{H}}$; ν_j choosen such that the power constraints in (4) are fulfilled with equality
2: **for** $k = 0, ..., K-1$ **do**
3: Transmit the data matrix $\mathbf{H}_j\mathbf{G}_j^{(k)}$ to other SCs;
4: Update the scaling factor $\beta^{(k)}$ according to (15)
5: Calculate the multiplier $\lambda_j^{(k)}$ using a root-finding algorithm to solve the equation (16);
6: Update the local precoding matrix $\mathbf{G}_j^{(k+1)}$ using (14);
7: **end for**

In addition, the Tx power constraint for the PR-DiP algorithm can be further extended to per-transmit antenna power constraint (PAPC), i.e., $\mathrm{E}\left\{\|\mathbf{g}_{nj}\|^2\right\} \leq P_{nj}$, where P_{nj} is the maximum Tx power for antenna n of SC j, and $\mathbf{g}_{nj} \in \mathbb{C}^{1\times N_{\mathrm{O}}}$ is the n-th row vector of precoding matrix $\mathbf{G}_j$. Thus, the update for $\mathbf{g}_{nj}$ under PAPC is straightforward according to the PR-DiP algorithm.

As proposed in [7] a variation of the PR-DiP algorithm, where the transmit vectors $\mathbf{x}_j$ instead of the precoding matrices $\mathbf{G}_j$ are updated, is also straightforward by using $\mathbf{x}_j = \mathbf{G}_j\mathbf{s}$. By updating the transmit signals $\mathbf{x}_j$, only the vectors $\mathbf{H}_j\mathbf{x}_j^{(k)} \in \mathbb{C}^{N_{\mathrm{O}}\times 1}$ instead of the matrices $\mathbf{H}_j\mathbf{G}_j^{(k)} \in \mathbb{C}^{N_{\mathrm{O}}\times N_{\mathrm{O}}}$ have to be exchanged at each iteration k. The drawback is, that the signals $\mathbf{x}_j$ have to be calculated for every time frame, while the precoding matrices can be kept constant as long as the channel does not change. A comparison of the total overhead for the update of the transmit signals $\mathbf{x}_j$ and the update of the precoding matrices $\mathbf{G}_j$ is given in [7].

IV. PERFORMANCE EVALUATION

In this section, we provide numerical simulations to evaluate the performance of the proposed distributed precoding algorithm. We present the averaged mean square error (aMSE) and the sum rate over all UEs for the considered algorithm simulated in an indoor hotspot (InH) dense scenario from 3GPP [22], where $N_{\rm SC} = 5$ SCs with $N_{\rm T} = 4$ Tx antennas are used to transmit signals to $N_{\rm UE} = 10$ UEs with $N_{\rm R} = 1$ Rx antenna. All SCs are deployed on the same floor, with a constant distance of 50 m between two of them, and are assumed to be connected with ideal links. Those UEs u are randomly dropped, each in a distance $d_{j,u}$ to SCs j on the same floor. For the channel matrix, each element is i.i.d. complex Gaussian distributed with zero mean and unit variance. Additionally, line-of-sight (LoS) transmission is assumed with path-loss from SC j to UE u defined as $PL_{j,u} = 16.9 \log_{10} d_{j,u} + 32.8 + 20 \log_{10} f_{\rm c}$, where the carrier frequency is $f_{\rm c} = 3.5$ GHz [23]. Each SC j is required to transmit with a normalized maximum power $P_j = 0.2$.

A. MSE Performance

To evaluate the performance of the proposed PR-DiP algorithm, we first define the metric of aMSE on the recovered signals:

$$\text{aMSE} = \frac{1}{N_{\rm UE}} \sum_{u=1}^{N_{\rm UE}} \text{E}\left\{ \|\mathbf{s}_u - \tilde{\mathbf{s}}_u\|^2 \right\} \tag{17}$$

As a benchmark, we show the aMSE performance of the centralized precoder under SPC which has a closed form analytic solution [8] and is given by

$$\mathbf{G} = \beta \left(\mathbf{H}^{\rm H}\mathbf{H} + \frac{N_0 \sigma_n^2}{P} \mathbf{I}_{N_{\rm T}} \right)^{-1} \mathbf{H}^{\rm H} = \beta \mathbf{G}'$$
$$\beta = \sqrt{\frac{P}{\text{tr}\left\{\mathbf{G}'\mathbf{G}'^{\rm H}\right\}}}, \tag{18}$$

where $P = \sum_{j=1}^{N_{\rm SC}} P_j = 1$ is the total transmit power. However the PSPC is a stricter constraint than the SPC, the centralized precoder under SPC shall serve as a benchmark for the proposed PR-DiP algorithm under PSPC because of its closed form solution.

In Fig. 2 the aMSE performance of our proposed PR-DiP algorithm under PSPC and the centralized solution under SPC given by (18) is shown w.r.t. the number of iterations for two different values of $\frac{\bar{E}_b}{N_0}$, i.e., the ratio between average received bit-energy and noise energy. For both $\frac{\bar{E}_b}{N_0}$, the PR-DiP algorithm converges to a solution which is slightly worse than the SPC solution due to the stricter power constraint. It can be observed, that for $\frac{\bar{E}_b}{N_0} = 10\,\text{dB}$ the algorithm converges faster than for $\frac{\bar{E}_b}{N_0} = 30\,\text{dB}$, which is also illustrated in Fig. 3. For $\frac{\bar{E}_b}{N_0} = 10\,\text{dB}$ the aMSE after $k = 20$ and $k = 50$ iterations are nearly equal, i.e., the algorithm after $k = 20$ iterations at $\frac{\bar{E}_b}{N_0} = 10\,\text{dB}$ is almost converged, while for higher $\frac{\bar{E}_b}{N_0}$ this gap increases. That is because, the noise power σ_n^2 is reflected in the update of the scaling factor $\beta^{(k)}$ in equation (15) and is therefore part of the diagonal term in the matrix to be inverted in (11). For very low SNR this diagonal term is dominating ${\beta^{(k)}}^2 \mathbf{\Lambda}^{(k)} \gg \mathbf{H}^{\rm H}\mathbf{H}$ and the matrix inversion in (11) can be approximated by

$$\left(\mathbf{H}^{\rm H}\mathbf{H} + {\beta^{(k)}}^2 \mathbf{\Lambda}^{(k)}\right)^{-1} \approx \left({\beta^{(k)}}^2 \mathbf{\Lambda}^{(k)}\right)^{-1} \approx \mathbf{T}. \tag{19}$$

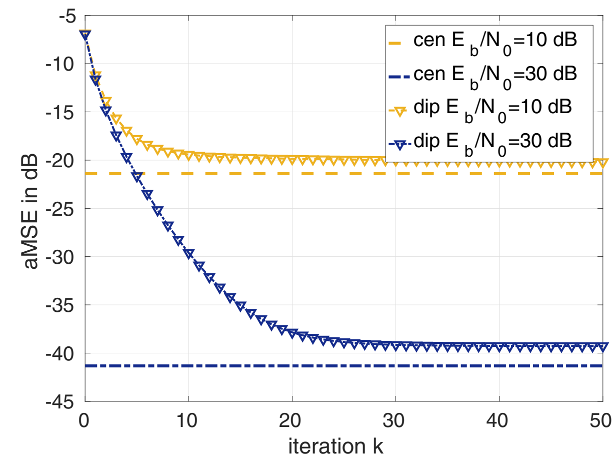

Fig. 2. aMSE w.r.t. no. of iterations for PR-DiP algorithm under PSPC and centralized precoder under SPC for $\frac{\bar{E}_b}{N_0} = 10\,\text{dB}$ and $\frac{\bar{E}_b}{N_0} = 30\,\text{dB}$

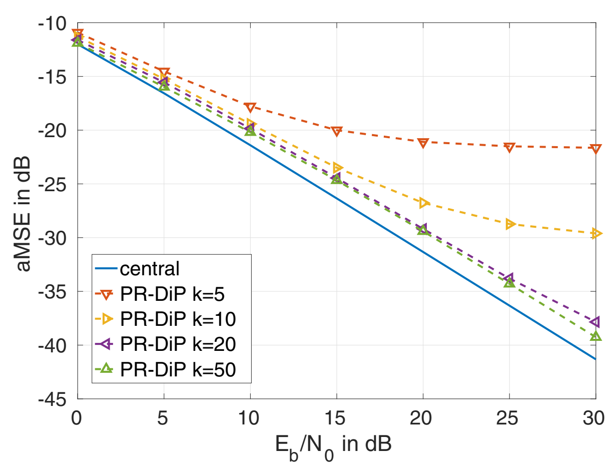

Fig. 3. aMSE w.r.t. $\frac{\bar{E}_b}{N_0}$ for proposed algorithm under different no. of iterations and centralized precoder

B. Partially Coordinated Transmission

Our PR-DiP algorithm can be easily modified for partially coordinated transmission, by setting the uth row of $\mathbf{H}_j$ and the uth column of $\mathbf{G}_j$ to zero, if UE u is not served by SC j. Then, two SCs only have to exchange the informations about the UEs served by both of them.

To show the gain due to cooperation between the SCs, it is assumed that each UE u is served by $N_{SC,u}$ SCs. In Fig. 4

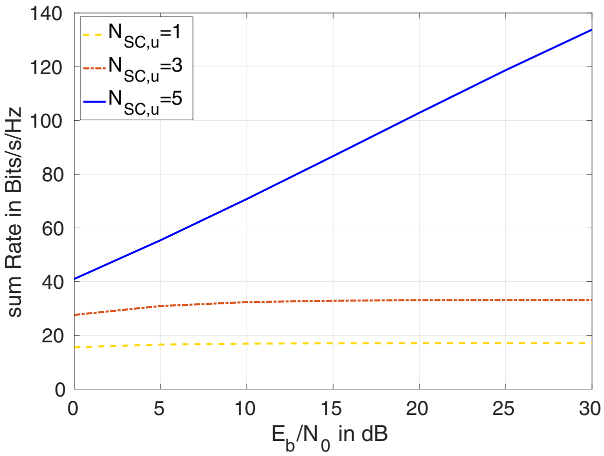

Fig. 4. Sum rate w.r.t. $\frac{\bar{E}_b}{N_0}$ for proposed algorithm using $K = 30$ iterations, considering no cooperation, partial cooperation and full cooperation

the sum rate $R = \sum_{u=1}^{N_{\text{UE}}} \log_2(1 + \text{SINR}_u)$ is shown w.r.t. $\frac{\bar{E}_b}{N_0}$ for different numbers of $N_{SC,u}$. Note that for $N_{SC,u} = 1$ there is no cooperation between the SCs and for $N_{SC,u} = 5$ all SCs j transmit its complete data matrix $\mathbf{H}_j \mathbf{G}_j^k$ to all other SCs. It can be observed, that there is a clear gap in the sum rate between complete joint transmission and partially coordinated transmission and this gap further increases with $\frac{\bar{E}_b}{N_0}$. Note, that only a small system with pure LoS transmission is considered here. Due to space limitations, an analysis of our proposed algorithm for large networks with mixed LoS and non-LoS transmission is left here.

V. Conclusion

In this paper, we provided a DiP algorithm for joint transmission in SC networks with PSPC. Starting from a constrained MMSE formulation, we rewrite the optimization problem into a form which could be numerically solved in a distributed way using the PR method. Simulations have shown promising results and further analysis of the overhead and computational complexity w.r.t. practical limitations, e.g., latency constraints and limited inter-SC link capacity, are planned for future study.

References

[1] S. Chen and J. Zhao, "The requirements, challenges, and technologies for 5g of terrestrial mobile telecommunication," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 36–43, May 2014.

[2] X. Ge, S. Tu, G. Mao, C. X. Wang, and T. Han, "5G Ultra-Dense Cellular Networks," *IEEE Wireless Communications*, vol. 23, no. 1, pp. 72–79, Feb. 2016.

[3] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," *IEEE Vehicular Technology Magazine*, vol. 6, no. 1, pp. 37–43, March 2011.

[4] P. Marsch and G. P. Fettweis, *Coordinated Multi-Point in Mobile Communications: From Theory to Practice*. Cambridge University Press, 2011.

[5] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell mimo cooperative networks: A new look at interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, December 2010.

[6] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 148–155, February 2012.

[7] G. Xu, H. Paul, T. Schier, P. Svedman, and A. Dekorsy, "Distributed precoding by in-network processing for small cell networks," in *European Wireless 2017 (EW17)*, May 2017.

[8] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2700–2712, Aug 2005.

[9] H. Paul, J. Fliege, and A. Dekorsy, "In-Network-Processing: Distributed Consensus-Based Linear Estimation," *IEEE Communications Letters*, vol. 17, no. 1, pp. 59–62, Jan. 2013.

[10] Y. Zou, C. Li, and L. Yang, "Distributed precoding for multicell MIMO networks," in *Wireless Communications and Signal Processing (WCSP), 2010 International Conference on*, Oct 2010.

[11] E. Björnson and B. Ottersten, "On the Principles of Multicell Precoding with Centralized and Distributed Cooperation," in *International Conference on Wireless Communications and Signal Processing, (WCSP 2009)*, Nanjing, China, 2009.

[12] R. Zhang, "Cooperative multi-cell block diagonalization with per-base-station power constraints," in *2010 IEEE Wireless Communication and Networking Conference*, April 2010.

[13] T. E. Bogale, L. Vandendorpe, and B. K. Chalise, "MMSE Transceiver design for coordinated base station systems: Distributive algorithm," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*, Nov 2010, pp. 1222–1226.

[14] Z. Kong, S. Yang, F. Wu, S. Peng, L. Zhong, and L. Hanzo, "Iterative Distributed Minimum Total MSE Approach for Secure Communications in MIMO Interference Channels," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 3, pp. 594–608, March 2016.

[15] H. Park, S. H. Park, H. B. Kong, and I. Lee, "Weighted Sum MSE Minimization under Per-BS Power Constraint for Network MIMO Systems," *IEEE Communications Letters*, vol. 16, no. 3, pp. 360–363, March 2012.

[16] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1989.

[17] Y. Ugur, Z. H. Awan, and A. Sezgin, "Cloud radio access networks with coded caching," in *WSA 2016; 20th International ITG Workshop on Smart Antennas*, March 2016.

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

[19] C. Windpassinger, *Detection and Precodingfor Multiple Input Multiple Output Channels*. Ph.D. dissertation, Uni. Erlangen-Nurnberg, Erlangen, Germany, 2004.

[20] S. Bazzi and W. Xu, "Robust Bayesian Precoding for Mitigation of TDD Hardware Calibration Errors," *IEEE Signal Processing Letters*, vol. 23, no. 7, pp. 929–933, July 2016.

[21] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed. New York, USA: Springer Science+Business Media, 2006.

[22] 3GPP TS 36.872, "Technical Specification Group Radio Access Network; Small cell enhancements for E-UTRA and E-UTRAN - Physical layer aspects (Release 12)," Dec. 2013.

[23] 3GPP TS 36.814, "Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9)," Mar. 2010.