Joint Decompression and Decoding for Cloud Radio Access Networks

Seok-Hwan Park, Osvaldo Simeone, Onur Sahin, and Shlomo Shamai (Shitz)

***Abstract*—In this work, joint decompression and decoding is studied for the uplink of multi-antenna cloud radio access networks. In this system, a set of multi-antenna mobile stations (MSs) wish to communicate with a "cloud" decoder through a set of multi-antenna base stations (BSs), which are connected to the cloud decoder through digital backhaul links of limited capacity. The BSs compress the received signal and send it to the cloud decoder, which performs joint decoding of the signals from all MSs. While the conventional solution prescribes that the cloud decoder performs first decompression and then decoding, recent work has shown that potentially larger rates can be achieved with joint decompression and decoding (JDD) at the cloud decoder. The sum-rate maximization problem with JDD, under the assumption of Gaussian test channels, is shown here to be an instance of a class of non-convex problems known as Difference of Convex (DC) problems. Based on this observation, an iterative algorithm based on the Majorization Minimization (MM) approach is proposed that guarantees convergence to a stationary point of the sum-rate maximization problem. Numerical results demonstrate the advantage of the proposed algorithm compared to the conventional approach based on separate decompression and decoding.**

***Index Terms*—Cloud radio access networks, distributed source coding, multi-cell processing, noisy network coding.**

I. INTRODUCTION

CLOUD radio access networks are by now recognized to provide a promising approach to solve "bandwidth crunch" problem and also to minimize the cost required for initial deployment or management of access points [1], [2]. On the uplink of the cloud radio access network, the base stations (BSs) operate as soft relay by compressing and forwarding the received signals to a cloud decoder through capacity-constrained backhaul links, as illustrated in Fig. 1.

Since the received signals at the different BSs are statistically correlated, it is beneficial to adopt distributed source coding strategies, as explored in [3]–[7]. In [4], a block-coordinate ascent algorithm was proposed to optimize the compression strategies at the BSs (via the corresponding test channels) under sum-backhaul capacity constraint. The work [6] instead considered

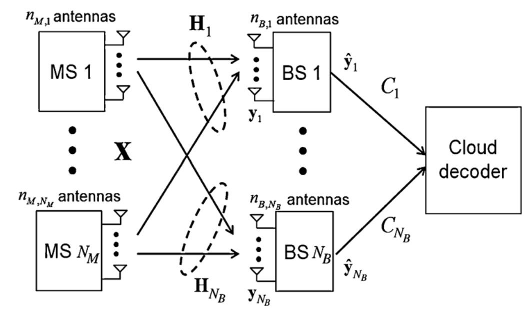

Fig. 1. Uplink of a cloud radio access network.

individual backhaul constraints and assumed sequential quantization with side information. In these previous works, the central decoder first decompresses the signals compressed by the BSs, and then decodes the mobile stations' (MSs) signals.

In this work, we consider a potentially more advantageous approach for the design of the cloud decoder, whereby the cloud decoder performs joint decompression and decoding (JDD). The idea was introduced and studied in [8], where specific results were given for single-antenna MSs and BSs. The goal of this work is to address the optimization of the compression strategies for *multi-antenna* BSs in the presence of JDD and *multi-antenna* MSs. To this end, we show that the sum-rate maximization problem, under the assumption of Gaussian test channels, is an instance of a class of non-convex problems known as Difference of Convex (DC) problems. Based on this observation, an iterative algorithm based on the Majorization Minimization (MM) approach is proposed that solves a sequence of convex problems obtained by linearizing the convex parts in the objective function of the original problem (see, e.g., [9]). It is shown that the proposed algorithm converges to a stationary point of the sum-rate maximization problem. The MM and related approaches were studied for beamforming matrix design in multi-cell downlink systems in [10] and for two-way relay channel models in [11]. From numerical results, we examine the advantage of the proposed JDD-based scheme over the conventional separate approach.

Notation: We use $p(y|x)$ to denote conditional probability density function (pdf) of random variable X given Y. All logarithms are in base two unless specified. Given vectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$, we define $\mathbf{x}_\mathcal{S}$ for a subset $\mathcal{S} \subseteq \{1, \ldots, m\}$ as the vector including, in ascending order, the vectors $\mathbf{x}_i$ with $i \in \mathcal{S}$; we set $\mathbf{x}_\phi$ as the empty vector. Similarly, given matrices $\mathbf{X}_1, \ldots, \mathbf{X}_m$, we denote by $\mathbf{X}_\mathcal{S}$ the matrix obtained by stacking the matrices $\mathbf{X}_i$ with $i \in \mathcal{S}$ vertically in ascending order. Notation $\mathbf{\Sigma_x}$ is used for the correlation matrix of random vector $\mathbf{x}$, i.e., $\mathbf{\Sigma_x} = \mathrm{E}[\mathbf{xx}^\dagger]$; and $\mathbf{\Sigma_{x|y}}$ represents the conditional correlation matrix of $\mathbf{x}$ given $\mathbf{y}$. We denote by $\mathrm{diag}(\{\mathbf{A}_j\}_{j\in\mathcal{S}})$ a block diagonal matrix whose diagonal blocks consists of the matrices $\mathbf{A}_j$ for $j \in \mathcal{S}$.

Manuscript received December 14, 2012; revised February 28, 2013; accepted March 08, 2013. Date of publication March 15, 2013; date of current version March 25, 2013. This research was supported by InterDigital, Inc. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chao Tian.

S.-H. Park and O. Simeone are with the Center for Wireless Communications and Signal Processing Research (CWCSPR), ECE Department, New Jersey Institute of Technology (NJIT), Newark, NJ 07102 USA (e-mail: seok-hwan.park@njit.edu; osvaldo.simeone@njit.edu).

O. Sahin is with InterDigital, Inc., Melville, NY 11747 USA (e-mail: Onur.Sahin@interdigital.com).

S. Shamai (Shitz) is with the Department of Electrical Engineering, The Technion, Haifa 32000, Israel (e-mail: sshlomo@ee.technion.ac.il).

Digital Object Identifier 10.1109/LSP.2013.2253095

II. System Model

We consider a *cluster* of cells, which includes a total number N_B of BSs and N_M active MSs. We denote the set of all BSs as $\mathcal{N}_B = \{1, \ldots, N_B\}$. Each ith BS is connected to the cloud decoder via a finite-capacity link of capacity C_i and has $n_{B,i}$ antennas, while each MS has $n_{M,i}$ antennas. Throughout the paper, we focus on the uplink, as illustrated in Fig. 1.

The overall channel from all MSs towards BS i is given as the $n_{B,i} \times n_M$ matrix

$$\mathbf{H}_i = [\mathbf{H}_{i1} \cdots \mathbf{H}_{iN_M}], \tag{1}$$

with $n_M = \sum_{i=1}^{N_M} n_{M,i}$ where we define $\mathbf{H}_{ij}$ as the $n_{B,i} \times n_{M,j}$ channel matrix between the jth MS and the ith BS. The signal received by the ith BS at a specific channel use (c.u.) is given by

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{x} + \mathbf{z}_i, \tag{2}$$

where vector $\mathbf{x} = [\mathbf{x}_1^\dagger \cdots \mathbf{x}_{N_M}^\dagger]^\dagger$ is the $n_M \times 1$ vector of symbols collecting all the $n_{M,j} \times 1$ vectors $\mathbf{x}_j$ transmitted by all MSs j, with $j = 1, \ldots, N_M$. The noise vectors $\mathbf{z}_i$ are independent across BS index $i \in \mathcal{N}_\mathcal{B}$ and have independent identically distributed (i.i.d.) entries with $\mathbf{z}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, for $i \in \mathcal{N}_\mathcal{B}$. We assume that the channel matrix $\mathbf{H}_i$ remains constant within each time-slot and is perfectly known by the cloud decoder. The signals $\mathbf{x}_j$ transmitted by each jth MS are assumed to be independent across index j and distributed as $\mathbf{x}_j \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{x}_j})$ for given correlation matrices $\mathbf{\Sigma}_{\mathbf{x}_j}$, $j = 1, \ldots, N_M$. It follows that we have $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_\mathbf{x})$ with $\mathbf{\Sigma}_\mathbf{x} = \mathrm{diag}(\mathbf{\Sigma}_{\mathbf{x}_1}, \ldots, \mathbf{\Sigma}_{\mathbf{x}_{N_M}})$.

Each jth BS communicates with the cloud by providing the latter a compressed version $\hat{\mathbf{y}}_j$ of the received signal $\mathbf{y}_j$. Note that this does not require the BSs to know the codebooks used by the MSs (but only their distribution). Using conventional information-theoretic arguments, a compression strategy for the jth BS is described by a test channel $p(\hat{\mathbf{y}}_j|\mathbf{y}_j)$ (see, e.g., [12]). Due to backhaul limitation, the compression at the jth BS is limited to C_j bits per c.u.. We are interested in designing the compression test channels $p(\hat{\mathbf{y}}_j|\mathbf{y}_j)$, with $j \in \mathcal{N}_\mathcal{B}$, at the BSs with the aim of maximizing the achievable sum-rate R_{sum}. The optimization is performed at the cloud decoder, which then informs the BSs about the optimal test channels.

III. Separate Decompression and Decoding

In this section, we review the sum-rate $R_{\mathrm{sum,SDD}}$ achievable with conventional separate decompression/decoding (SDD) approach at the cloud decoder [4], [6], [8]. Accordingly, the cloud decoder first decompresses the signals $\hat{\mathbf{y}}_j$ and then, based on all signals $\hat{\mathbf{y}}_{\mathcal{N}_\mathcal{B}} = [\hat{\mathbf{y}}_1, \ldots, \hat{\mathbf{y}}_{N_B}]$, decodes the MSs' messages. For decompression, fix an ordering π of the BS indices. The cloud decoder decompresses in the order $\hat{\mathbf{y}}_{\pi(1)}, \hat{\mathbf{y}}_{\pi(2)}, \ldots, \hat{\mathbf{y}}_{\pi(N_B)}$. Therefore, when decompressing $\hat{\mathbf{y}}_{\pi(i)}$, the cloud decoder has already retrieved the signals $\hat{\mathbf{y}}_{\pi(1)}, \ldots, \hat{\mathbf{y}}_{\pi(i-1)}$. These can be treated as side information available at the decoder but not to the encoder, namely BS $\pi(i)$. As a result, using arguments similar to the Wyner-Ziv theorem [13], the description $\hat{\mathbf{y}}_{\pi(i)}$ for $i \in \mathcal{N}_\mathcal{B}$ can be recovered at the cloud decoder if the test channels $p(\hat{\mathbf{y}}_i|\mathbf{y}_i)$ and the backhaul capacities C_i for $i \in \mathcal{N}_\mathcal{B}$ satisfy the conditions

$$I\left(\mathbf{y}_{\pi(i)}; \hat{\mathbf{y}}_{\pi(i)} | \hat{\mathbf{y}}_{\{\pi(1),\ldots,\pi(i-1)\}}\right) \le C_{\pi(i)}. \tag{3}$$

As mentioned, the cloud decodes jointly the signals $\mathbf{x}$ of all MSs based on all the descriptions $\hat{\mathbf{y}}_i$ for $i \in \mathcal{N}_\mathcal{B}$, so that the achievable sum-rate for given test channels $\{p(\hat{\mathbf{y}}_i|\mathbf{y}_i)\}_{i\in\mathcal{N}_\mathcal{B}}$ is given by

$$R_{\mathrm{sum,SDD}} = I\left(\mathbf{x}; \hat{\mathbf{y}}_{\mathcal{N}_\mathcal{B}}\right), \tag{4}$$

and the sum-rate maximization problem with SDD is stated as

$$\begin{aligned} \underset{\pi, \{p(\hat{\mathbf{y}}_i|\mathbf{y}_i)\}_{i\in\mathcal{N}_\mathcal{B}}}{\text{maximize}} \quad & R_{\mathrm{sum,SDD}} \\ \text{s.t.} \quad & I\left(\mathbf{y}_{\pi(i)}; \hat{\mathbf{y}}_{\pi(i)} | \hat{\mathbf{y}}_{\{\pi(1),\ldots,\pi(i-1)\}}\right) \le C_{\pi(i)} \\ & \text{for all } i \in \mathcal{N}_\mathcal{B}. \end{aligned} \tag{5}$$

Note that the optimization space includes the test channels $\{p(\hat{\mathbf{y}}_i|\mathbf{y}_i)\}_{i\in\mathcal{N}_\mathcal{B}}$ as well as the BS ordering π. In Section V, we will consider two ordering methods, exhaustive search, which requires a search over all possible orderings and greedy ordering, which successively chooses the best BS corresponding to the largest rate increase (see [6] for more detail). We refer to [3]–[6] for further details on the solution of problem (5).

IV. Joint Decompression and Decoding

With conventional SDD, the central processor in the cloud decoder recovers the compressed signals $\hat{\mathbf{y}}_{\mathcal{N}_B}$ first, and then performs joint decoding of all of the MSs' signals $\mathbf{x}$. In this section, instead, we assume that the cloud decoder performs joint decompression and decoding (JDD), i.e., joint decompression of the signals $\hat{\mathbf{y}}_{\mathcal{N}_B}$ and decoding of the signals $\mathbf{x}$. It is remarked that errors in decompression do not affect the system performance as long as the signals $\mathbf{x}$ are correctly decoded [8]. The sum-rate achievable with the JDD strategy, denoted by $R_{\mathrm{sum,JDD}}$, was derived in [8], [14] for a generic channel model, as summarized in Lemma 1 below. Our main contribution is to propose an algorithm for the optimization of the test channels in the cloud radio access model under study.

Lemma 1: [8], [14]: For given test channels $\{p(\hat{\mathbf{y}}_j|\mathbf{y}_j)\}_{j\in\mathcal{N}_\mathcal{B}}$, the following sum-rate $R_{\mathrm{sum,JDD}}$ is achievable with the JDD.

$$R_{\mathrm{sum,JDD}} = \min_{\mathcal{S}\subseteq\mathcal{N}_\mathcal{B}} \left\{ \sum_{j\in\mathcal{S}} \left(C_j - I(\mathbf{y}_j; \hat{\mathbf{y}}_j|\mathbf{x})\right) + I(\mathbf{x}; \hat{\mathbf{y}}_{\mathcal{S}^c}) \right\}. \tag{6}$$

To interpret the sum-rate (6), we observe that for each subset $\mathcal{S}$ of BSs, the rate $\sum_{j\in\mathcal{S}}(C_j - I(\mathbf{y}_j; \hat{\mathbf{y}}_j|\mathbf{x})) + I(\mathbf{x}; \hat{\mathbf{y}}_{\mathcal{S}^c})$ equals the sum of: *i*) the overall backhaul capacity $\sum_{j\in\mathcal{S}} C_j$ for the BSs in set $\mathcal{S}$ discounted by the total rate wasted in compressing quantization noise, namely $\sum_{j\in\mathcal{S}} I(\mathbf{y}_j; \hat{\mathbf{y}}_j|\mathbf{x})$; *ii*) the sum-rate $I(\mathbf{x}; \hat{\mathbf{y}}_{\mathcal{S}^c})$ that would be achievable based only on the signals received by the BSs in $\mathcal{S}^c$. Achievability of (6) is proved in [8], [14].

Using Lemma 1, the sum-rate maximization problem with the JDD is formulated as

$$\underset{\{p(\hat{\mathbf{y}}_j|\mathbf{y}_j)\}_{j\in\mathcal{N}_\mathcal{B}}}{\text{maximize}} \quad R_{\mathrm{sum,JDD}}. \tag{7}$$

Without claim of optimality, we adopt Gaussian test channels, namely $\hat{\mathbf{y}}_j = \mathbf{y}_j + \mathbf{q}_j$ with $\mathbf{q}_j \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Omega}_j)$, for $j \in \mathcal{N}_\mathcal{B}$, which is known to be optimal with SDD [3]–[6]. With this choice, the mutual information measures appearing in (7) are computed as

$$I(\mathbf{y}_j;\hat{\mathbf{y}}_j|\mathbf{x}) = \log|\mathbf{I}+\mathbf{A}_j|, \tag{8}$$

$$\text{and } I(\mathbf{x};\hat{\mathbf{y}}_{\mathcal{S}^c}) = \log\left|\mathbf{I}+\mathbf{\Sigma}_{\mathbf{y}_{\mathcal{S}^c}}\operatorname{diag}\left(\{\mathbf{A}_j\}_{j\in\mathcal{S}^c}\right)\right| - \sum_{j\in\mathcal{S}^c}\log|\mathbf{I}+\mathbf{A}_j|, \tag{9}$$

where we defined $\mathbf{A}_j = \mathbf{\Omega}_j^{-1}$ and the covariance $\mathbf{\Sigma}_{\mathbf{y}_{\mathcal{S}^c}}$ is computed as $\mathbf{\Sigma}_{\mathbf{y}_{\mathcal{S}^c}} = \mathbf{I}+\mathbf{H}_{\mathcal{S}^c}\mathbf{\Sigma}_{\mathbf{x}}\mathbf{H}_{\mathcal{S}^c}^{\dagger}$. Moreover, the problem (7) becomes

$$\underset{\{\mathbf{A}_j\succeq\mathbf{0}\}_{j\in\mathcal{N}_\mathcal{B}}}{\text{maximize}}\ \min_{\mathcal{S}\subseteq\mathcal{N}_\mathcal{B}}\left\{\sum_{j\in\mathcal{S}}C_j - \sum_{j\in\mathcal{N}_\mathcal{B}}\log|\mathbf{I}+\mathbf{A}_j| + \log\left|\mathbf{I}+\mathbf{\Sigma}_{\mathbf{y}_{\mathcal{S}^c}}\operatorname{diag}\left(\{\mathbf{A}_j\}_{j\in\mathcal{S}^c}\right)\right|\right\}. \tag{10}$$

Since the objective function in (10) is not smooth, it is convenient to reformulate (10) by considering the epigraph form [15, Sec. 4.1.3] of (10) as

$$\begin{aligned}\underset{\{\mathbf{A}_j\succeq\mathbf{0}\}_{j\in\mathcal{N}_\mathcal{B}},\gamma}{\text{maximize}}\ &\gamma - \sum_{j\in\mathcal{N}_\mathcal{B}}\log|\mathbf{I}+\mathbf{A}_j|\\ \text{s.t. } &\gamma - \log\left|\mathbf{I}+\mathbf{\Sigma}_{\mathbf{y}_{\mathcal{S}^c}}\operatorname{diag}\left(\{\mathbf{A}_j\}_{j\in\mathcal{S}^c}\right)\right| - \sum_{j\in\mathcal{S}}C_j \le 0,\ \text{for all } \mathcal{S}\subseteq\mathcal{N}_\mathcal{B}.\end{aligned} \tag{11}$$

The problem (11) is not convex due to the second term in the objective function which is a non-linear convex, and thus not concave, function of $\mathbf{A}_j$ (the constraints are instead convex functions as desired). However, the objective function in (11) is the difference of two convex functions. Therefore, problem (11) is a so called DC problem. For this class of problems, various algorithms are known that have desirable properties [9]–[11]. Here we consider the so called MM approach [9], which solves a sequence of convex problems (see (12)) obtained by linearizing the convex parts in the objective function. The resulting MM algorithm, summarized in Algorithm 1, provides a sequence of achievable rates $R_{\text{sum,JDD}}^{(n)}$ for each iteration n, whose properties are given in the Lemma 2.

Algorithm 1 MM Algorithm for problem (11)

1. Initialize the matrix $\mathbf{A}_j^{(1)}$ to an arbitrary positive semidefinite matrix for $j\in\mathcal{N}_\mathcal{B}$ and set $n=1$.
2. Update the matrices $\mathbf{A}_j^{(n+1)}$ for $j\in\mathcal{N}_\mathcal{B}$ as a solution of the following (convex) problem.

$$\begin{aligned}\underset{\left\{\mathbf{A}_j^{(n+1)}\succeq\mathbf{0}\right\}_{j\in\mathcal{N}_\mathcal{B}},\gamma}{\text{maximize}}\ &\gamma - \sum_{j\in\mathcal{N}_\mathcal{B}} g\left(\mathbf{A}_j^{(n+1)},\mathbf{A}_j^{(n)}\right)\\ \text{s.t. } &\gamma - \log\left|\mathbf{I}+\mathbf{\Sigma}_{\mathbf{y}_{\mathcal{S}^c}}\operatorname{diag}\left(\left\{\mathbf{A}_j^{(n+1)}\right\}_{j\in\mathcal{S}^c}\right)\right| - \sum_{j\in\mathcal{S}}C_j \le 0,\ \text{for all } \mathcal{S}\subseteq\mathcal{N}_\mathcal{B},\end{aligned} \tag{12}$$

where the function $g(\mathbf{A}_j^{(n+1)},\mathbf{A}_j^{(n)})$ is a linear function of $\mathbf{A}_j^{(n+1)}$ defined as

$$g\left(\mathbf{A}_j^{(n+1)},\mathbf{A}_j^{(n)}\right) = \log\left|\mathbf{I}+\mathbf{A}_j^{(n)}\right| + \frac{1}{\ln 2}\operatorname{tr}\left(\left(\mathbf{I}+\mathbf{A}_j^{(n)}\right)^{-1}\left(\mathbf{A}_j^{(n+1)}-\mathbf{A}_j^{(n)}\right)\right). \tag{13}$$

3. Stop if $\sum_{j\in\mathcal{N}_\mathcal{B}}\|\mathbf{A}_j^{(n+1)}-\mathbf{A}_j^{(n)}\|_F^2 < \delta$ with predefined threshold value δ. Otherwise, set $n\leftarrow n+1$ and go back to Step 2.

Lemma 2: The sequence $R_{\text{sum,JDD}}^{(n)}$ is monotonically increasing with respect to iteration index n, and it converges to a stationary point of the problem (11) as $n\rightarrow\infty$.

Proof: The proof follows the same steps as [10, Theorem 1] and is thus omitted. ∎

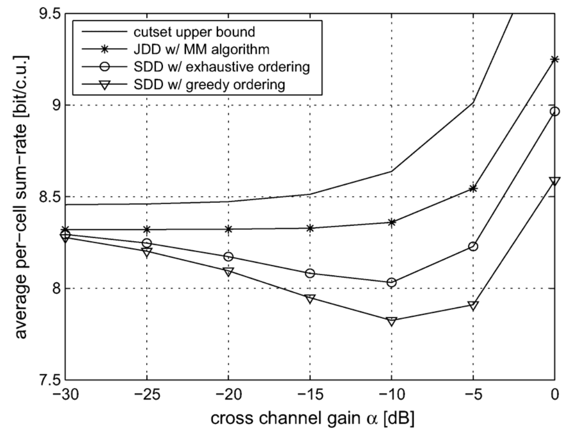

Fig. 2. Average per-cell sum-rate versus the inter-cell channel gain α with $N_B=3, n_{M,i}=1, n_{B,i}=4, C_i=12$ bit/c.u. and $P=20$ dB.

V. Numerical Results

In this section, we present numerical results to investigate the advantage of the proposed JDD based scheme as compared to the conventional SDD schemes. For simplicity, it is assumed that all the MSs use a single transmit antenna with equal transmit power P which leads to the covariance $\Sigma_{\mathbf{x}} = P\mathbf{I}$. Moreover, we assume $N_B = N_M$ and there is one MS active in each cell. The elements of the channel matrix $\mathbf{H}_{i,k}$ between the MS in the kth cell and the BS in the ith cell are i.i.d. complex Gaussian distributed with $\mathcal{CN}(0,\alpha^{|i-k|})$. We fix the number of cells to three, i.e., $N_B = N_M = 3$. The achievable rate is averaged over the realization of the channel matrices.

In Fig. 2, we plot the average per-cell rate (i.e., sum-rate divided by N_B) versus the inter-cell channel gain α with $n_{B,i} = 4$, $C_i = 12$ bit/c.u. and $P = 20$ dB. We consider two ordering methods for SDD, namely exhaustive search and greedy ordering (see Section III). For reference, we also plot the cutset upper bound [16, Theorem 14.10.1] R_{cutset}, which is computed as

$$R_{\text{cutset}} = \min_{\mathcal{S}\subseteq\mathcal{N}_\mathcal{B}}\left\{\sum_{j\in\mathcal{S}}C_j + \log\left|\mathbf{I}+\mathbf{H}_{\mathcal{S}^c}\mathbf{\Sigma}_{\mathbf{x}}\mathbf{H}_{\mathcal{S}^c}^{\dagger}\right|\right\}. \tag{14}$$

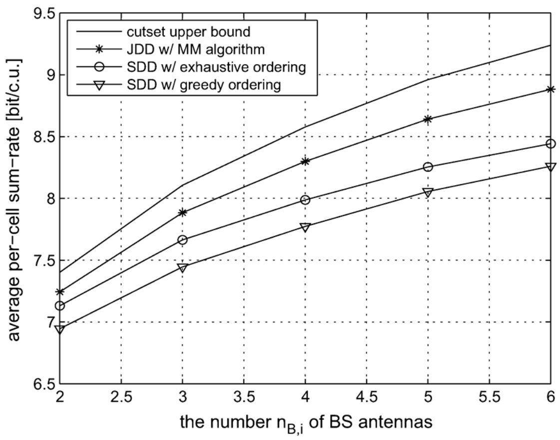

Fig. 3. Average per-cell sum-rate versus the number $n_{B,i}$ of BS antennas with $N_B = 3, n_{M,i} = 1, \alpha = -10$ dB, $C_i = 12$ bit/c.u. and $P = 20$ dB.

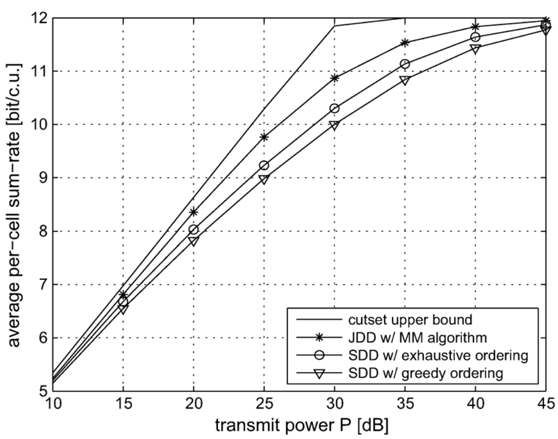

Fig. 4. Average per-cell sum-rate versus the transmit power P_k per user with $N_B = 3, n_{B,i} = 4, n_{M,i} = 1, \alpha = -10$ dB, and $C_i = 12$ bit/c.u..

From Fig. 2, it is observed that an increase in the inter-cell channel gain α affects the performance of the SDD schemes in two conflicting ways: *i*) it increases the overall system signal-to-noise ratio which has a positive impact; *ii*) it increases the compression rate (3) required to keep a given compression fidelity on the backhaul. As a result, for sufficiently large α, the first effect dominates and the sum-rate of SDD increases, while, for lower values of α, the second effect is more relevant and the sum-rate of SDD decreases. In contrast, JDD always benefits from the increased channel power since the backhaul penalty term $I(\mathbf{y}_j; \hat{\mathbf{y}}_j|\mathbf{x})$ in (6) is not affected by the channel matrices or transmit signals' powers, but it depends only on the quantization noise as seen in (8).

Fig. 3 demonstrates the impact of the number $n_{B,i}$ of BS antennas when $\alpha = -10$ dB, $C_i = 12$ bit/c.u. and $P = 20$ dB. It is seen that the gain of the JDD scheme is more pronounced when the BSs have a larger number of antennas. This is because, as the received signals lie in a large dimensional spaces, more sophisticated and efficient compression strategies are called for. Finally, we plot the per-user average sum-rate versus the transmit power P in Fig. 4 with $n_{B,i} = 4$, $\alpha = -10$ dB and $C_i = 12$ bit/c.u.. The JDD scheme shows relevant rate gains in the regime of moderate-to-large P in which the rate is not limited by the capacity of the backhaul links. Moreover, from all the figures, it can be observed that JDD performs quite close to the cutset upper bound (14).

VI. Conclusions

In this work, we tackled the problem of optimizing the compression strategies at the BSs for the uplink of a cloud radio access network. We aimed at maximizing the sum-rate achievable with joint decompression of the signals received by the BSs and decoding of the MSs' messages at the cloud decoder. The proposed iterative solution solves a sequence of convex problems and produces a sequence of feasible points with increasing sum-rate that converges to a stationary point of the problem. Numerical results showed the advantage of the proposed scheme compared to the conventional SDD-based schemes. An open problem is the management of the complexity of the algorithm for systems with a large number of BSs, e.g., via appropriate cell clustering techniques.

References

[1] P. Marsch, B. Raaf, A. Szufarska, P. Mogensen, H. Guan, M. Farber, S. Redana, K. Pedersen, and T. Kolding, "Future mobile communication networks: Challenges in the design and operation," *IEEE Veh. Technol. Mag.*, vol. 7, no. 1, pp. 16–23, Mar. 2012.
[2] J. Segel and M. Weldon, "Lightradio portfolio-technical overview," Technology White Paper 1, Alcatel-Lucent.
[3] G. Chechik, A. Globerson, N. Tishby, and Y. Weiss, "Information bottleneck for Gaussian variables," *J. Mach. Learn. Res.*, vol. 6, pp. 165–188, 2005.
[4] A. del Coso and S. Simoens, "Distributed compression for MIMO coordinated networks with a backhaul constraint," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4698–4709, Sep. 2009.
[5] C. Tian and J. Chen, "Remote vector Gaussian source coding with decoder side information," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4676–4680, Oct. 2009.
[6] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai (Shitz), "Robust and efficient distributed compression for cloud radio access networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 692–703, Feb. 2013.
[7] L. Zhou and W. Yu, "Uplink multicell processing with limited backhaul via successive interference cancellation," in *Proc. IEEE Glob. Comm. Conf. (Globecom 2012)*, Anaheim, CA, Dec. 2012.
[8] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai (Shitz), "Uplink macro diversity of limited backhaul cellular network," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3457–3478, Aug. 2009.
[9] A. Beck and M. Teboulle, "Gradient-based algorithms with applications to signal recovery problems," in *Convex Optimization in Signal Processing and Communications*, Y. Eldar and D. Palomar, Eds. Cambridge, U.K.: Cambridge University Press, 2010, pp. 42–88.
[10] M. Hong, Q. Li, Y.-F. Liu, and Z.-Q. Luo, "Decomposition by Successive Convex Approximation: A Unifying Approach for Linear Transceiver Design in Interfering Heterogeneous Networks," [Online]. Available: arXiv:1210.1507
[11] A. Khabbazibasmenj, F. Roemer, S. A. Vorobyov, and M. Haardt, "Sum-rate maximization in two-way AF MIMO relaying: Polynomial time solutions to a class of DC programming problems," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5478–5493, Oct. 2010.
[12] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[13] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.
[14] S. H. Lim, Y.-H. Kim, A. E. Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.
[15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[16] T. Cover and J. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecomm., 1st ed. New York, NY, USA: Wiley, 1991.