Transmit Beamformer and Quantization Design for Multi-Carrier CRAN CoMP

Ganesh Venkatraman, Antti Tölli, Jarkko Kaleva, Markku Juntti

Centre for Wireless Communications (CWC) - Radio Technologies,

University of Oulu, Finland - FI-90014, P.O. Box 4500, Email: firstname.lastname@ee.oulu.fi

Abstract—We consider the problem of designing transmit beamformers for a multi-carrier joint processing coordinated multi-point (CoMP) transmission with finite backhaul capacity. We assume a Cloud Radio access network (CRAN) wherein the beamformers are designed centrally, quantized and sent via the backhaul to the respective base stations (BSs). The beamformer design becomes a nonconvex combinatorial problem with the user selection, which is also modeled as a part of our problem. In order to avoid the combinatorial search, we relax the binary association variable by a linear one and the nonconvex sets by a sequence of convex subsets by the successive convex approximation (SCA) method. Thereby, we design the beamformers by iteratively solving a convex subproblem in each SCA iteration. Additionally, to enforce sparsity in the relaxed binary variable, we regularize the objective with the associated entropy measure to enforce a binary solution. Since the beamformers are notified via backhaul to the respective BSs, we include the beamformer overhead also in the backhaul usage in addition to data sharing. Numerical results are provided to illustrate the performance of the proposed designs.

I. Introduction

We consider a cooperative multi-cell joint processing coordinated multi-point (CoMP) transmission wherein multiple base stations (BSs) transmit the same data to users in a coordinated manner over orthogonal frequency division multiplexing (OFDM) framework [1]. Serving each user by multiple BSs can increase the respective user throughput. However, due to the finite backhaul capacity, it is often not possible to include all BSs in the network for CoMP transmission as it demands the knowledge of users data at multiple BSs. Furthermore, as the number of users increases, CoMP joint processing (JP) requires huge signaling via backhaul. Therefore, it is not always beneficial to share the data of all users in the system to all BSs. In such cases, the beamformer design becomes a more challenging problem as it involves both user association and the beamformer design for the chosen subset of users [2]–[4].

Downlink transmission is primarily focused on satisfying certain quality-of-service (QoS) imposed by the higher layers based on the user service category. These QoS constraints are related to emptying the associated user queued packets with some priorities in the physical layer. Therefore, most of the beamformer designs are aimed at minimizing the number of backlogged packets. Earlier studies on queue minimization were summarized in the survey papers [5]. In particular, the problem of power allocation to minimize the number of backlogged packets was considered in [6] using geometric programming. Since the beamformers are designed by a centralized controller (CC), they have to be signaled to the respective BSs via the finite backhaul link, therefore, to reduce the signaling overhead, only the quantized beamformers are fed back to the BSs. Beamformer quantization for multiple-input single-output (MISO) transmissions has been studied extensively in the literature [7]–[11].

II. System Model & Problem Formulation

We consider a model with N_B BSs, each equipped with N_T transmit antenna elements, serving K single antenna users over N OFDM sub-channels. The set of all BS indices available in the

system is denoted by $\mathcal{B} = \{1, 2, \ldots, N_B\}$ and the set of all users is represented by $\mathcal{U} = \{1, 2, \ldots, K\}$. Now, the received symbol $y_{k,n}$ of user k on the nth sub-channel is given by

$$y_{k,n} = \sum_{b \in \mathcal{B}} \sum_{k \in \mathcal{U}} \mathbf{h}_{b,k,n} \mathbf{m}_{b,k,n} d_{k,n} + e_{k,n} \tag{1}$$

where $\mathbf{h}_{b,k,n} \in \mathbb{C}^{1 \times N_T}$ denotes the channel between user k and BS b on the nth OFDM sub-channel. The transmit data symbol $d_{k,n}$ corresponding to each user k has $\mathbb{E}[|d_{k,n}|^2] = 1$ and $e_{k,n}$ is a sample of circularly symmetric additive white complex Gaussian noise drawn from $\mathcal{CN}(0, N_0)$. The transmit beamformers $\mathbf{m}_{b,k,n} \in \mathbb{C}^{N_T \times 1}$, used to transmit data to user k, are designed by the CC and provided to the respective BSs in $\mathcal{B}$ for coordinated transmission. The received signal-to-interference-plus-noise ratio (SINR) of user k is given as

$$\gamma_{k,n} = \frac{\left| \sum_{b \in \mathcal{B}} \mathbf{h}_{b,k,n} \mathbf{m}_{b,k,n} \right|^2}{\sum_{k' \in \mathcal{U} \setminus k} \left| \sum_{b \in \mathcal{B}} \mathbf{h}_{b,k,n} \mathbf{m}_{b,k',n} \right|^2 + N_0} \tag{2}$$

Let Q_k be the number of backlogged packets destined for user k at a given scheduling instant. The queue dynamics of user k are modeled using the Poisson arrival process with the average number of packet arrivals of $\Lambda_k = \mathbf{E}_i\{\lambda_k\}$ packets or bits, where $\lambda_k[i] \sim \mathrm{Pois}(\Lambda_k)$ represents the instantaneous number of packets arriving for user k at the ith time instant. The total number of queued packets at the $(i + 1)$th instant for user k, denoted as $Q_k[i + 1]$, is given by

$$Q_k[i + 1] = \max \{Q_k[i] - t_k[i], 0\} + \lambda_k[i] \tag{3}$$

where t_k denotes the number of transmitted packets or bits for user k over N OFDM sub-channels. The maximum transmission rate achieved in the ith slot is given by $t_k[i] \leq \sum_{n=1}^{N} \log_2(1 + \gamma_{k,n}[i])$ for a SINR $\gamma_{k,n}$. Note that t_k and Q_k are represented by the same units, *i.e.*, in bits defined per channel use. For brevity, we drop the time index i in the forthcoming discussions.

In order to minimize the number of backlogged packets associated with the users, we consider CoMP model with the objective

$$\underset{t_{k,n}}{\mathrm{minimize}} \quad \sum_{k \in \mathcal{U}} a_k \left| Q_k - \sum_{n=1}^{N} t_{k,n} \right|_q \tag{4}$$

where $t_{k,n}$ is the rate seen by user k on the nth sub-channel and a_k is a constant which can be used to emphasize the user priority. The exponent $q \in \mathbb{Z}_{++}$ determines how the backlogged packets of the users are to be treated [12]. The objective (4) not only minimizes the number of backlogged packets, but it also reduces the average packet delay by setting a_k as Λ_k^{-1} along with $q = 1$.

We assume that the channel state information (CSI) knowledge is available at the CC for designing transmit beamformers as it involves BSs to CC traffic. Even though the CSI has finite resolution, it can be improved by differential feedback due to time-correlated nature of the channels. However, due to the instantaneous arrival process, transmit beamformers corresponding to the BSs need not be the same even if the associated channels are constant. Therefore, the backhaul link between the CC and BS is used to signal both data

and the beamformers. Due to the finite backhaul capacity, transmit beamformers are quantized before signaling them to the corresponding BSs. We assume that each complex entry of the beamformer vector is quantized with N_Q bits. Since multiple data symbols in a frame are precoded with a same beamformer, it is efficient to send beamformers and data symbols from finite alphabets separately to the respective BSs. However, if the BSs act as a remote radio head, then signaling the beamformers and data symbols separately is not an efficient approach. Hence, we consider the former scenario.

If we assume that BS b is serving user k, it is not guaranteed that user k is served by BS b on all the sub-channels. Therefore, to evaluate the data requirement via the backhaul to serve user k by BS b over all sub-channels, we introduce a binary variable $c_{b,k,n} \in \{0, 1\}$, which is assigned to one when BS b is serving (or aiding) user k on the nth sub-channel and zero otherwise. Thus, the total signaling via the finite capacity backhaul between BS b and the CC is given as

$$\sum_{n=1}^{N} \sum_{k \in \mathcal{U}} c_{b,k,n} t_{k,n} + N_Q N_T \sum_{n=1}^{N} \sum_{k \in \mathcal{U}} c_{b,k,n} \leq B_b, \forall b \in \mathcal{B} \quad (5)$$

where $N_Q N_T$ is the total number of bits required to signal each transmit beamformer and B_b is the finite backhaul capacity between BS b and the CC.

Now, the optimization variable $t_{k,n}$ in (4) is determined by the SINR $\gamma_{k,n}$, which in turn depends on the transmit beamformers by (2). In order to ensure that $c_{b,k,n}$ determines whether BS b is serving user k over a sub-channel n or not, the binary variable $c_{b,k,n}$ should be coupled with the respective beamformer as

$$\|\mathbf{m}_{b,k,n}\|^2 \leq c_{b,k,n} \nu_{b,k,n} \quad (6)$$

where $\nu_{b,k,n}$ is an additional optimization variable, which corresponds to the power allocated by BS b to user k on sub-channel n. Therefore, when $c_{b,k,n}$ is zero, the respective transmit beamformer $\mathbf{m}_{b,k,n}$ will be a zero vector, thereby ensuring the absence of transmission from BS b to user k on the nth sub-channel. Let $\mathcal{O} = \{0, 1, \ldots, N\}$ be the set of all sub-channel indices and $\{\mathbf{m}\} \triangleq \{\mathbf{m}_{b,k,n}\}, \forall b \in \mathcal{B}, \forall k \in \mathcal{U}, \forall n \in \mathcal{O}$ be the collection of all transmit beamformers. Similarly, $\{\mathbf{t}\}$ and $\{\boldsymbol{\nu}\}$ denotes the collection of all users rate from each sub-channel and the transmit power from all BSs to each user. Now, the problem of designing transmit beamformers is given as

$$\begin{aligned} \underset{\substack{\{\mathbf{t}\},\{\mathbf{m}\},\\ \{\boldsymbol{\gamma}\},\{\boldsymbol{\nu}\},\{\mathbf{c}\}}}{\text{minimize}} \quad & \sum_{k \in \mathcal{U}} a_k \Big| Q_k - \sum_{n=1}^{N} t_{k,n} \Big|_q & (7a) \\ \text{subject to} \quad & \log(1 + \gamma_{k,n}) \geq t_{k,n} & (7b) \\ & \|\mathbf{m}_{b,k,n}\|^2 \leq c_{b,k,n} \nu_{b,k,n} & (7c) \\ & \sum_{k \in \mathcal{U}} \sum_{n=1}^{N} \nu_{b,k,n} \leq P_{\max}, \forall b & (7d) \\ & c_{b,k,n} \in \{0, 1\}, \forall b, \forall k, \forall n & (7e) \\ & \sum_{n=1}^{N} \sum_{k \in \mathcal{U}} c_{b,k,n} (t_{k,n} + N_Q N_T) \leq B_b, \forall b & (7f) \end{aligned}$$

where $\gamma_{k,n}$ is defined in (2). The constraint (7d) is included to limit the total transmit power from each BS within $P_{\max}$. The constraint (7b) introduces $t_{k,n}$ as an under-estimator for the rate seen by user k on sub-channel n, since $t_{k,n} = \log(1 + \gamma_{k,n})$ cannot be handled directly in (7) as a convex constraint.

III. BEAMFORMER DESIGN WITH FIXED QUANTIZATION LEVEL

In order to solve (7), we relax the SINR expression in (2) as

$$\frac{\left| \sum_{b \in \mathcal{B}} \mathbf{h}_{b,k,n} \mathbf{m}_{b,k,n} \right|^2}{b_{k,n}} \geq \gamma_{k,n} \quad (8a)$$

$$\sum_{k' \in \mathcal{U} \setminus k} \left| \sum_{b \in \mathcal{B}} \mathbf{h}_{b,k,n} \mathbf{m}_{b,k',n} \right|^2 + N_0 \leq b_{k,n} \quad (8b)$$

since (2) cannot be handled directly as a convex constraint. The constraint (8a) is an under-estimator for $\gamma_{k,n}$ and $b_{k,n}$ in (8b) is a over-estimator for the total interference seen by user k on sub-channel n. Even after relaxing (2) with (8), the resulting problem is not convex due to the constraints (8a) and (7f). Therefore, we adopt successive convex approximation (SCA) method in [13], [14], where the nonconvex constraints are replaced by a sequence of convex subsets that can be solved iteratively until convergence.

To find a convex subset for each of the nonconvex constraint, let us begin with (8a) by considering an equivalent representation for the SINR constraint in (8a) related to user k on sub-channel n as

$$\gamma_{k,n} \leq \frac{|\bar{\mathbf{h}}_{k,n} \bar{\mathbf{m}}_{k,n}|^2}{b_{k,n}} \quad (9)$$

where the stacked channel vector $\bar{\mathbf{h}}_{k,n} \in \mathbb{C}^{1 \times |\mathcal{B}| N_T}$ and transmit precoder $\bar{\mathbf{m}}_{k,n} \in \mathbb{C}^{|\mathcal{B}| N_T \times 1}$ are defined as

$$\bar{\mathbf{h}}_{k,n} \triangleq \left[\mathbf{h}_{\mathcal{B}(1),k,n}, \mathbf{h}_{\mathcal{B}(2),k,n}, \ldots, \mathbf{h}_{\mathcal{B}(|\mathcal{B}|),k,n} \right] \quad (10a)$$

$$\bar{\mathbf{m}}_{k,n} \triangleq \left[\mathbf{m}_{\mathcal{B}(1),k,n}^{\mathrm{T}}, \mathbf{m}_{\mathcal{B}(2),k,n}^{\mathrm{T}}, \ldots, \mathbf{m}_{\mathcal{B}(|\mathcal{B}|),k,n}^{\mathrm{T}} \right]^{\mathrm{T}}. \quad (10b)$$

The fractional term in (9) is of *convex-over-linear* form, and thus can be bounded by the linear first order Taylor approximation as

$$\begin{aligned} \mathcal{L}_{k,n}^{(i)}(\bar{\mathbf{m}}_{k,n}, b_{k,n}; \bar{\mathbf{m}}_{k,n}^{(i)}, b_{k,n}^{(i)}) &\triangleq 2 \frac{\bar{\mathbf{m}}_{k,n}^{(i)} \bar{\mathbf{h}}_{k,n}^{\mathrm{H}} \bar{\mathbf{h}}_{k,n}}{b_{k,n}^{(i)}} \left(\bar{\mathbf{m}}_{k,n} - \bar{\mathbf{m}}_{k,n}^{(i)} \right) \\ &+ \frac{|\bar{\mathbf{h}}_{k,n} \bar{\mathbf{m}}_{k,n}^{(i)}|^2}{b_{k,n}^{(i)}} \left(1 - \frac{b_{k,n} - b_{k,n}^{(i)}}{b_{k,n}^{(i)}} \right) \leq \frac{|\bar{\mathbf{h}}_{k,n} \bar{\mathbf{m}}_{k,n}|^2}{b_{k,n}} \quad (11) \end{aligned}$$

where the linear function $\mathcal{L}_{k,n}^{(i)}(\bar{\mathbf{m}}_{k,n}, b_{k,n})$ is an under-estimator for the r.h.s of (9) with equality at the operating point $\{\bar{\mathbf{m}}_{k,n}^{(i)}, \bar{\mathbf{b}}_{k,n}^{(i)}\}$. Similarly, the linear approximation can be used to find a convex subset for the nonconvex constraint (7f) as

$$\begin{aligned} \mathcal{D}_{b}^{(i)}(c_{b,k,n}, t_{k,n}; c_{b,k,n}^{(i)}, t_{k,n}^{(i)}) &\triangleq \sum_{n=1}^{N} \sum_{k \in \mathcal{U}} \Big[c_{b,k,n}^{(i)} t_{k,n}^{(i)} + \\ c_{b,k,n}^{(i)} \left(t_{k,n} - t_{k,n}^{(i)} \right) + t_{k,n}^{(i)} \left(c_{b,k,n} - c_{b,k,n}^{(i)} \right) \Big] &\geq c_{b,k,n} t_{k,n} \quad (12) \end{aligned}$$

where the operating points $c_{b,k,n}^{(i)}$ and $t_{k,n}^{(i)}$ are the solutions for $c_{b,k,n}$ and $t_{k,n}$, respectively in the $(i-1)$th SCA iteration.

Now, by replacing (8a) with (11) and (7f) by (12), we obtain a mixed-integer quadratic programming (MIQP) subproblem in each step that can be solved with the existing solvers [15]–[17]. The fixed operating points $\{\mathbf{m}^{(i)}\} = \{\mathbf{m}_{b,k,n}^{(i)}\}, \forall b, \forall k, \forall n$ and $\{\mathbf{b}^{(i)}\} = \{b_{k,n}^{(i)}\}, \forall k, \forall n$ are updated with the solution from previous SCA iteration $i-1$, where the ith MIQP subproblem is given as

$$\begin{aligned} \underset{\substack{\{\mathbf{t}\},\{\mathbf{m}\},\{\mathbf{b}\},\\ \{\boldsymbol{\gamma}\},\{\boldsymbol{\nu}\},\{\mathbf{c}\}}}{\text{minimize}} \quad & \sum_{k \in \mathcal{U}} a_k \Big| Q_k - \sum_{n=1}^{N} t_{k,n} \Big|_q & (13a) \\ \text{subject to} \quad & \mathcal{L}_{k,n}^{(i)}(\bar{\mathbf{m}}_{k,n}, b_{k,n}; \bar{\mathbf{m}}_{k,n}^{(i)}, b_{k,n}^{(i)}) \geq \gamma_{k,n} & (13b) \\ & \mathcal{D}_{b}^{(i)}(c_{b,k,n}, t_{k,n}; c_{b,k,n}^{(i)}, t_{k,n}^{(i)}) \\ & \quad + \mu N_Q N_T \sum_{n=1}^{N} \sum_{k \in \mathcal{U}} c_{b,k,n} \leq B_b, \forall b & (13c) \\ & (7b) - (7e) \text{ and } (8b) & (13d) \end{aligned}$$

where μ in (13c) accounts for the precoder overhead in the backhaul utilization, i.e., the number of data symbols for which the precoder is used, which depends on the channel coherency.

Due to the complexity involved with (13), we propose a relaxation for the binary decision variable as $c_{b,k,n} \in [0,1]$. However, by doing so, the resulting solution for $c_{b,k,n}$ need not be binary. Therefore, to encourage binary solution for the linear variable $c_{b,k,n}$, we include an additional constraint in the formulation to enforce sparsity in the solution. One such function is the entropy measure, defined as

$$\mathcal{H}_b(\mathbf{c}_b) \triangleq -\sum_{k=1}^{K}\sum_{n=1}^{N} c_{b,k,n} \log(c_{b,k,n}) \qquad (14)$$

where $\mathbf{c}_b$ is a vector formed by stacking all $c_{b,k,n}, \forall k$ and $\forall n$ corresponding to BS b. In order to ensure binary outcome for $\mathbf{c}_b$, $\mathcal{H}_b(\mathbf{c}_b)$ in (14) must be equal to or less than zero. By doing so, the constraint with $\mathcal{H}_b(\mathbf{c}_b) \leq 0$ will become nonconvex due to the concavity of the entropy function. Thus, we relax the constraint by the partial Lagrangian approach with dual variables $\tau_b, \forall b \in \mathcal{B}$ as

$$\underset{\substack{\{\mathbf{t}\},\{\mathbf{m}\},\{\mathbf{b}\}, \\ \{\gamma\},\{\nu\},\{\mathbf{c}\}}}{\text{minimize}} \quad \sum_{k\in\mathcal{U}} a_k \Big| Q_k - \sum_{n=1}^{N} t_{k,n} \Big|_q + \sum_{b\in\mathcal{B}} \tau_b \, \mathcal{H}_b(\mathbf{c}_b) \qquad (15a)$$

$$\text{subject to} \quad 0 \leq c_{b,k,n} \leq 1, \ \forall b, \forall k, \forall n \qquad (15b)$$

$$(7b) - (7d), (8b), (13b), \text{ and } (13c) \qquad (15c)$$

which is similar to the reweighted ℓ_1 minimization method in [18].

The problem in (15) is nonconvex due to the concave nature of the entropy function. Therefore, to formulate a tractable problem, we employ the SCA technique to find a linear over-estimator for the entropy function as

$$\mathcal{H}_b(\mathbf{c}_b) \leq \widetilde{\mathcal{H}}_b(\mathbf{c}_b; \mathbf{c}_b^{(i)}) \triangleq -\sum_{k=1}^{K}\sum_{n=1}^{N} c_{b,k,n}^{(i)} \log\left(c_{b,k,n}^{(i)}\right) - \\ \sum_{k=1}^{K}\sum_{n=1}^{N}\left(1 + \log\left(c_{b,k,n}^{(i)}\right)\right)\left(c_{b,k,n} - c_{b,k,n}^{(i)}\right) \qquad (16)$$

which is a first order Taylor approximation of the entropy function. Now, for a fixed set of τ, the relaxed convex subproblem for the SCA iteration i is given as

$$\underset{\substack{\{\mathbf{t}\},\{\mathbf{m}\},\{\mathbf{b}\}, \\ \{\gamma\},\{\nu\},\{\mathbf{c}\}}}{\text{minimize}} \quad \sum_{k\in\mathcal{U}} a_k \Big| Q_k - \sum_{n=1}^{N} t_{k,n} \Big|_q + \sum_{b\in\mathcal{B}} \tau_b^{(i)} \, \widetilde{\mathcal{H}}_b(\mathbf{c}_b; \mathbf{c}_b^{(i)}) \ (17a)$$

$$\text{subject to} \quad 0 \leq c_{b,k,n} \leq 1, \ \forall b, \forall k, \forall n \qquad (17b)$$

$$(7b) - (7d), (8b), (13b), \text{ and } (13c) \qquad (17c)$$

where $\tau_b^{(i)}, \forall b$ are the fixed dual variables, given by the subgradient update as

$$\tau_b^{(i+1)} = \tau_b^{(i)} + \rho \left[\widetilde{\mathcal{H}}_b(\mathbf{c}_b^{(i+1)}; \mathbf{c}_b^{(i)})\right]^{+} \qquad (18)$$

where $\mathbf{c}_b^{(i+1)}$ is the solution for $\mathbf{c}_b$ in the ith iteration, ρ is a fixed step size with $\rho^{(i)} > 0$ to ensure sparse solution upon convergence and $[x]^{+} \triangleq \max(x,0)$. As $c_{b,k,n} \to \{0,1\}$, $\mathcal{H}_b(\mathbf{c}_b) \to 0$, thereby $\widetilde{\mathcal{H}}_b(\mathbf{c}_b; \mathbf{c}_b^{(i)})$ in (17a) has no effect on the objective of (17).

IV. BEAMFORMER DESIGN WITH VARIABLE QUANTIZATION

In this section, we study the impact of quantized beamformers in terms of the number of residual packets for a CoMP scenario. In order to do so, we introduce an additional error vector $\mathbf{e} \in \mathbb{C}^{N_T \times 1}$ in the transmit beamformer corresponding to the finite resolution error as $\hat{\mathbf{m}}_{b,k,n} = \mathbf{m}_{b,k,n} + \mathbf{e}_{b,k,n}$. Unlike the traditional approach wherein a single codebook is used, we assume that each entity has multiple codebooks corresponding to variable quantization levels. Therefore, signaling an entry in the codebook includes both the index in the quantization table and the entry corresponding to the table itself.

The quantization error $\mathbf{e}_{b,k,n}$ is defined for each beamformer and has the maximum distortion $\delta_{b,k,n}$, i.e., $\|\mathbf{e}_{b,k,n}\|^2 \leq \delta_{b,k,n}^2$, where $\delta_{b,k,n} \in \mathbb{R}_+$ is the maximum magnitude distortion that can occur in

the beamformer and it depends on the number of quantization bits N_Q, i.e., $\delta_{b,k,n} \leq 2^{-(N_Q+1)} A_{\max}$, assuming uniform quantizer for magnitude and the dynamic range of each beamformer entry is given by $A_{\max} = [0, \sqrt{P_{\max}}]$ [19]. For simplicity, we assume uniform quantization of both magnitude and phase of each beamformer entry $\mathbf{m}_{b,k,n}(i), \forall i \in \{1,2,\ldots,N_T\}$. Now, the SINR seen by user k with the quantized beamformer is given by

$$\tilde{\gamma}_{k,n} = \frac{\left|\sum_{b\in\mathcal{B}} \mathbf{h}_{b,k,n}\hat{\mathbf{m}}_{b,k,n}\right|^2}{\sum_{k'\in\mathcal{U}\backslash k}\left|\sum_{b\in\mathcal{B}} \mathbf{h}_{b,k,n}\hat{\mathbf{m}}_{b,k',n}\right|^2 + N_0}. \qquad (19)$$

Since neither the quantization level nor the transmit beamformers are known, it is not possible to use (19) instead of (2) in (17). Therefore, by using the worst case error model for (19), we have

$$\hat{\gamma}_{k,n} = \frac{\left|\sum_{b\in\mathcal{B}} \mathbf{h}_{b,k,n}\left(\mathbf{m}_{b,k,n} - \mathbf{g}_{b,k,n}^{(i)}\delta_{b,k,n}\right)\right|^2}{\sum_{k'\in\mathcal{U}\backslash k}\left|\sum_{b\in\mathcal{B}} \mathbf{h}_{b,k,n}\left(\mathbf{m}_{b,k',n} + \mathbf{g}_{b,k',n}^{(i)}\delta_{b,k',n}\right)\right|^2 + N_0} \qquad (20)$$

where $\mathbf{g}_{b,k,n}^{(i)} = \mathbf{m}_{b,k,n}^{(i)}/\|\mathbf{m}_{b,k,n}^{(i)}\|$ is a unit vector in the direction of the respective transmit beamformer obtained from the SCA iteration $(i-1)$th. Instead of using $\mathbf{m}_{b,k,n}^{(i)}$, we have used the previous beamformer $\mathbf{m}_{b,k,n}^{(i)}$ for $\mathbf{g}_{b,k,n}^{(i)}$ to obtain a tractable problem. As the SCA iteration $i \to \infty$, the gap between the current and the previous beamformers diminishes, i.e., $\|\mathbf{m}_{b,k,n} - \mathbf{m}_{b,k,n}^{(i)}\| \to 0$.

Following the same approach as in Section III, we can find a convex approximation for the worst case SINR expression in the presence of quantization error as in (11) wherein $\mathbf{m}_{b,k,n}$ is replaced by the quantized version, i.e., $\mathbf{m}_{b,k,n} - \mathbf{g}_{b,k,n}^{(i)}\delta_{b,k,n}$, in (10b). Similarly, the interference term $b_{k,n}$ in (8b) can be given as

$$\sum_{k'\in\mathcal{U}\backslash k}\left|\sum_{b\in\mathcal{B}} \mathbf{h}_{b,k,n}\left(\mathbf{m}_{b,k',n} + \mathbf{g}_{b,k',n}^{(i)}\delta_{b,k',n}\right)\right|^2 + N_0 \leq b_{k,n} \ (21)$$

and the objective function is given by (17a). Unlike problem (17), the number of quantization bits N_Q is not fixed as it is determined by $\delta_{b,k,n}$. Now, the constraint in (5) can be written as

$$\sum_{n=1}^{N}\sum_{k\in\mathcal{U}} c_{b,k,n}t_{k,n} + \mu' N_T \sum_{n=1}^{N}\sum_{k\in\mathcal{U}} c_{b,k,n} \log_2\left(\frac{A_{\max}}{2\delta_{b,k,n}}\right) \leq B_b \qquad (22)$$

which follows from the uniform quantizer assumption. As it is evident that (22) is a nonconvex constraint, we resort to the SCA technique to find a convex subset for (22) by linearizing it with the first order Taylor approximation as

$$\sum_{n=1}^{N}\sum_{k\in\mathcal{U}}\left[c_{b,k,n}^{(i)}t_{k,n}^{(i)} + c_{b,k,n}^{(i)}\left(t_{k,n} - t_{k,n}^{(i)}\right)\right.$$
$$+ t_{k,n}^{(i)}\left(c_{b,k,n} - c_{b,k,n}^{(i)}\right) + \mu' N_T \, c_{b,k,n}^{(i)} \log_2\left(\frac{A_{\max}}{2}\right)$$
$$-\mu' N_T \log_2(e)\left(c_{b,k,n}^{(i)} \log\left(\delta_{b,k,n}^{(i)}\right) + \log\left(\delta_{b,k,n}^{(i)}\right)\left(c_{b,k,n} - c_{b,k,n}^{(i)}\right)\right.$$
$$\left.\left. + \frac{c_{b,k,n}^{(i)}}{\delta_{b,k,n}^{(i)}}\left(\delta_{b,k,n} - \delta_{b,k,n}^{(i)}\right)\right)\right] \leq B_b. \quad (23)$$

In order to improve the SINR in (20), $\delta_{b,k,n}$ should be minimized. However, by doing so, $-\log(\delta_{b,k,n})$ increases, thereby reducing the available backhaul capacity for the data symbols. Therefore, $\delta_{b,k,n}$ should be minimized only for those users who can achieve noticeable SINR improvement. However, for the unserved users, $\delta_{b,k,n}$ can take zero without increasing the backhaul utilization, since $c_{b,k,n} = 0$. It

is important that $\delta_{b,k,n} = 0$ for all links with $c_{b,k,n} = 0$ in order to eliminate $\delta_{b,k,n}$ from (20) for all $c_{b,k,n} = 0$. On the contrary, if $P_{\max}$ is large, $\delta_{b,k,n}$ can take unity and force $-\log(\delta_{b,k,n}) = 0$ without affecting the backhaul utilization. In order to avoid such issues, we regularize the objective with the quantization error $\delta_{b,k,n}$ with a constant $\psi > 0$. The convex subproblem for the ith SCA iteration is given by

$$
\underset{\substack{\{\mathbf{t}\},\{\mathbf{m}\},\{\mathbf{b}\},\\\{\boldsymbol{\delta}\}\{\boldsymbol{\gamma}\},\{\boldsymbol{\nu}\},\{\mathbf{c}\}}}{\text{minimize}} \quad \sum_{k\in\mathcal{U}} a_k \Big| Q_k - \sum_{n=1}^{N} t_{k,n} \Big|_q + \sum_{b\in\mathcal{B}} \tau_b^{(i)} \, \widetilde{\mathcal{H}}_b(\mathbf{c}_b; \mathbf{c}_b^{(i)})
$$

$$
+ \psi \sum_{b\in\mathcal{B}} \sum_{n=1}^{N} \sum_{k=1}^{K} \delta_{b,k,n} \quad \text{(24a)}
$$

$$
\text{subject to} \quad 0 \le c_{b,k,n} \le 1, \quad \delta_{b,k,n} \ge 0 \quad \text{(24b)}
$$

$$
\sum_{k=1}^{K}\sum_{n=1}^{N}\left(\nu_{b,k,n} + \delta_{b,k,n}^2\right) \le P_{\max}, \forall b \quad \text{(24c)}
$$

$$
(7b), (7c), (13b), (21) \text{ and } (23) \quad \text{(24d)}
$$

where (24c) includes both beamformer power and the variance of $\delta_{b,k,n}$ in the power budget. It follows from the uncorrelated nature of the quantization error to the beamformers. The overhead involved in specifying the codebook index can be captured in μ' in (22).

V. Numerical Results

We demonstrate the performance of the proposed CoMP beamformer designs in terms of the average number of backlogged bits. The path loss seen by the users are drawn uniformly from $[-3,0]$ dB and the channels are drawn from $h_{b,k,n} \sim \mathcal{CN}(0;1)$. The noise variance is fixed as $N_0 = 1$ in all simulations. In Fig. 1, we denote coordinated transmission by *static* approach wherein each user is served by a BS based on the path loss. In addition, we also compared with the *beam selection* model, labeled in Fig. 1, which is formed by including an additional constraint $\sum_{b\in\mathcal{B}} c_{b,k,n} = 1, \forall k \forall n$ to (17), i.e., on any sub-channel, a user can be served by a single BS only.

While considering beamformer quantization in Fig. 1, the backhaul is used to notify both the user data and the beamformers. However, for schemes without quantization, the backhaul is used only to notify the user data and not the beamformers as it is assumed to be known perfectly at the BSs. For variable quantization model (24), we use ten different codebooks with $N_Q = \{3, \ldots, 12\}$ bits respectively, and the N_Q bits are divided equally among the phase and amplitude of each complex entry. Depending on distortion $\delta_{b,k,n}$, a codebook is chosen as in Section IV to quantize the beamformers. Using quantized beamformers, the actual SINR and user rate are evaluated.

Fig. 1(a) plots the total number of residual packets (bits) in the system for various schemes by varying the average packet arrival rates of the users. It can be seen that the centralized unquantized CoMP has the least number of residual packets, followed by the *beam selection* scheme and then by the *static* scenario. However, in the presence of beamformer quantization, the total number of backlogged packets increases significantly for all schemes invariably. Fig. 1(a) shows that the proposed beamformer design based on worst-case model for CoMP has a noticeable reduction in the number of backlogged packets compared to other schemes. Increasing the quantization level improves the CoMP performance marginally. Due to the impact of quantization error on the coherent transmission from multiple BSs, the quantized *beam selection* is superior to the CoMP design.

Fig. 1(b) illustrates the average number of BSs serving each active user in the system for various values of average packet arrivals. The number of BSs serving each active user is maximum in the CoMP model when the total arrival bits is less than the backhaul capacity. However, as the arrival rate increases, we can see from Fig. 1(b) that the number of serving BSs decreases to ≈ 2. On the contrary,
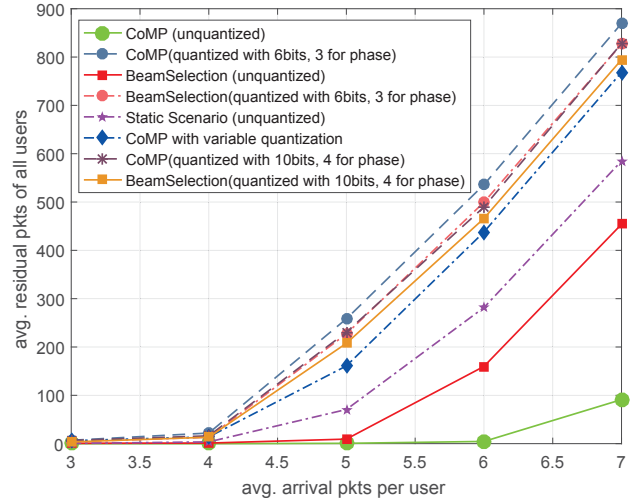
(a). Average Number of backlogged packets.
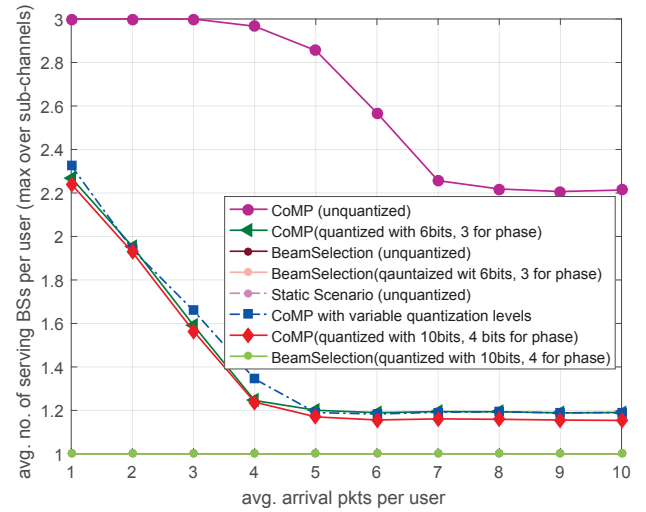
(b). Number of serving BSs per user.

Fig. 1. Behavior of model $\{N_B, K, N, N_T\} = \{3, 6, 2, 4\}$ over 100 slots with backhaul capacity of $[20, 30, 40]$ and $\mu = 0.1$

including precoder quantization as well in the backhaul utilization, the number serving BSs is less than the one without considering the precoder quantization. As the arrival rate increases, Fig. 1(b) shows that the CoMP model with precoder quantization converges to a single BS transmission model, which is similar to *beam selection* scheme.

VI. Conclusions

We proposed two related transmit beamformer designs for a multi-carrier coordinated multi-point transmission with finite backhaul capacity. Due to the nonconvexity and the combinatorial nature of the proposed problems, we resorted to the successive convex approximation technique by relaxing the nonconvex sets by a sequence of convex subsets and relaxed the binary constraints by the linear ones. Moreover, to obtain binary solution for the relaxed variables, we regularized the objective with the entropy measure of the relaxed variables. Since the beamformers are signaled via backhaul along with the data, we proposed a beamformer design to alter each quantization level based on the worst-case model. Numerical results were provided to illustrate the robustness of the proposed beamformer design in the presence of quantization errors.

REFERENCES

[1] 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 J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, 2010.

[2] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Downlink multicell processing with limited-backhaul capacity," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, pp. 1–10, 2009.

[3] B. Dai and W. Yu, "Sparse beamforming and user-centric clustering for downlink cloud radio access network," *IEEE Access*, vol. 2, pp. 1326–1339, 2014.

[4] J. Zhao, T. Q. Quek, and Z. Lei, "Coordinated multipoint transmission with limited backhaul data transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2762–2775, 2013.

[5] M. Chiang, S. Low, A. Calderbank, and J. Doyle, "Layering as Optimization Decomposition: A Mathematical Theory of Network Architectures," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255–312, Jan 2007.

[6] K. Seong, R. Narasimhan, and J. Cioffi, "Queue Proportional Scheduling via Geometric Programming in Fading Broadcast Channels," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1593–1602, 2006.

[7] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct 1998.

[8] S. Bhashyam, A. Sabharwal, and B. Aazhang, "Feedback gain in multiple antenna systems," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 785–798, May 2002.

[9] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, 2003.

[10] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, 2003.

[11] D. J. Love, R. W. Heath Jr, W. Santipach, and M. L. Honig, "What is the value of limited feedback for mimo channels?" *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 54–59, 2004.

[12] G. Venkatraman, A. Tölli, M. Juntti, and L.-N. Tran, "Traffic Aware Resource Allocation Schemes for Multi-Cell MIMO-OFDM Systems," *IEEE Trans. Signal Process.*, vol. 64, pp. 2730–2745, Jun. 2016.

[13] B. R. Marks and G. P. Wright, "A General Inner Approximation Algorithm for Nonconvex Mathematical Programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, 1978.

[14] G. Scutari, F. Facchinei, L. Lampariello, and P. Song, "Distributed Methods for Constrained Nonconvex Multi-Agent Optimization – Part I: Theory." [Online]. Available: http://arxiv.org/abs/1410.4754v1

[15] J. Löfberg, "YALMIP : A Toolbox for Modeling and Optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. [Online]. Available: http://users.isy.liu.se/johanl/yalmip

[16] Gurobi Optimization, Inc., "Gurobi Optimizer Reference Manual," 2015. [Online]. Available: http://www.gurobi.com

[17] MOSEK ApS, *The MOSEK Optimization Toolbox for MATLAB Manual. Version 7.1 (Revision 28).*, 2015. [Online]. Available: http://docs.mosek.com/7.1/toolbox/index.html

[18] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing Sparsity by Reweighted ℓ_1 Minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[19] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct 1998.