Group Sparse Precoding for Cloud-RAN with Multiple User Antennas

Zhiyang Liu · Hong Wu · Yingxin Zhao

Abstract Cloud radio access network (C-RAN) has become a promising network architecture to support the massive data traffic in the next generation cellular networks. In a C-RAN, a massive number of low-cost remote antenna ports (RAPs) are connected to a single baseband unit (BBU) pool via high-speed low-latency fronthaul links, which enables efficient resource allocation and interference management. As the RAPs are geographically distributed, the group sparse beamforming schemes attract extensive studies, where a subset of RAPs is assigned to be active and a high spectral efficiency can be achieved. However, most studies assumes that each user is equipped with a single antenna. How to design the group sparse precoder for the multiple antenna users remains little understood, as it requires the joint optimization of the mutual coupling transmit and receive beamformers. This paper formulates an optimal joint RAP selection and precoding design problem in a C-RAN with multiple antennas at each user. Specifically, we assume a fixed transmit power constraint for each RAP, and investigate the optimal tradeoff between the sum rate and the number of active RAPs. Motivated by the compressive sensing theory, this paper formulates the group sparse precoding problem by inducing the $\ell_0$-norm as a penalty and then uses the reweighted $\ell_1$ heuristic to find a solution. By adopting the idea of block diagonalization precoding, the problem can be formulated as a convex optimization, and an efficient algorithm is proposed based on its Lagrangian dual. Simulation results verify that our proposed algorithm can achieve almost the same sum rate as that obtained from exhaustive search.

Keywords cloud radio access network · sparse beamforming · block diagonalization · group-sparsity · antenna selection
1 Introduction

Cloud radio access network (C-RAN) [1] is a promising and flexible architecture to accommodate the exponential growth of mobile data traffic in the next-generation cellular network. In a C-RAN, all the base-band signal processing is shifted to a single base-band unit (BBU) pool [2]. In the meantime, the conventional base-stations (BSs) are replaced by geographically distributed remote antenna ports (RAPs) with only antenna elements and power amplifiers, which are connected to the BBU pool via high-speed low-latency fronthaul links by fiber [3]. The simple structure of RAP enables the probability to deploy ultra-dense RAPs in a C-RAN with low cost.

With highly densed geographically distributed RAPs, significant rate gains can be expected over that with the same amount of co-located antennas in both the single-user and multi-user cases [4–6]. Due to the huge differences among the distance between the user and the geographically distributed RAPs, it has been shown in [4] that the capacity in the single-user case is crucially determined by the access distance from the user to its closest RAP. This motivates us to investigate whether it is possible to achieve a significant proportion of the sum rate by using a subset of the RAPs. In the single-user case, it is straightforward to avoid distant RAPs transmitting, as they have little contribution to improve the capacity. In the multi-user case, the problem becomes challenging, as the beamformers of all users should be jointly designed.

To tackle this problem, a branch of sparse beamforming technologies are therefore proposed, where the beamforming vectors are designed to be sparse with respect to the total number of transmit antennas [7–9]. Motivated by the recent theoretical breakthroughs in compressive sensing [10], the sparse beamforming problem is formulated by including the $\ell_1$ norm of the beamforming vectors as a regularization such that the problem becomes convex. By iteratively updating the weights of the $\ell_1$ norm, the sparse beamformer that minimizes the total transmit power can be obtained by iteratively solving a second-order-conic-programming (SOCP) [7] or a semi-definite-programming (SDP) [8]. The problem can be further simplified to a uplink beamformer design problem via uplink-downlink duality [9]. Nevertheless, when each RAP includes multiple antennas, one RAP will be switched off only when all the coefficients in its beamformer are set to be zero. In other word, all antennas at a RAP should be selected or ignored simultaneously, which requires group sparsity instead of individual sparsity as in conventional compressive sensing. Recently, group sparse beamforming problem are proposed which can be formulated by inducing a mixed $\ell_1/\ell_p$ norm regularization [11–13]. For instance, by inducing the $\ell_1/\ell_2$-norm as a regularization, the weighted sum rate maximization problem is formulated as a weighted minimum mean square error (WMMSE) minimization and can be solved via a quadratical-constrained-quadratic-programming (QCQP) [11]. Compared to the individual sparse beamforming, the group sparse beamforming further reduces the network power consumption, and the energy efficiency can be improved as well.

So far, most algorithms focus on the situation where each user has a single antenna [7–9,11–13]. As suggested by the multiple-input-multiple-output (MIMO) theory, the capacity increases linearly with the minimum number of transmit and receive antennas [14]. In a C-RAN, it is desirable to employ multiple antennas at each user to exploit the potential multiplexing gains, which, however, further
complicates the sparse precoder design. The difficulty originates from the fact that the problem is typically non-convex, and the transmit and receive beamformers, which are mutual coupling, should be jointly designed. This paper focuses on designing group sparse precoder based on block diagonalization (BD) [15], which has gained widespread popularity thanks to its low complexity and near-capacity performance when the number of transmit antennas is large [16–19].

In this paper, we address the joint problem of RAP selection and joint precoder design in a C-RAN with multiple antennas at each user and each RAP. Whereas the problem is typically NP-hard, we show that the problem becomes convex by inducing the reweighted $\ell_1$ norm of a vector that indicates the transmit power at each RAP as a regularization. Based on its Lagrangian dual problem, we propose an iterative algorithm by iteratively updating the weights of the $\ell_1$ norm to generate sparse solution. Simulation results verify that the proposed algorithm can achieve almost the same sum rate as that achieved from exhaustive search.

The rest of this paper is organized as follows: Section II introduces the system model and formulates the problem. Section III proposes an iterative algorithm to solve the group sparse precoding problem. Simulation results are presented and discussed in Section IV. Section V concludes this paper.

Throughout this paper, italic letters denote scalars, and boldface upper-case and lower-case letters denote matrices and vectors, respectively. The superscripts $T$ and $\dagger$ denote transpose and conjugate transpose, respectively. $\mathbb{E}\{\cdot\}$ denotes the expectation operator. $\lceil \cdot \rceil$ denote the ceiling operator. $\|x\|_p$ denotes the $\ell_p$ norm of vector $x$. $\text{Tr}\{X\}$ and $\text{det}\{X\}$ denote the trace and determinant of matrix $X$, respectively. $\text{diag}(a_1, \ldots, a_N)$ denotes an $N \times N$ diagonal matrix with diagonal entries $\{a_i\}$. $I_N$ denotes an $N \times N$ identity matrix. $0_{N \times M}$ and $1_{N \times M}$ denote $N \times M$ matrices with all entries zero and one, respectively. $|X|$ denotes the cardinality of set $X$.

2 System Model and Problem Formulation

Consider a C-RAN with a set of remote antenna ports (RAPs), denoted as $\mathcal{L}$, and a set of users, denoted as $\mathcal{K}$, with $|\mathcal{L}| = L$ and $|\mathcal{K}| = K$, as shown in Fig. 1. Suppose that each RAP is equipped with $N_c$ antennas, and each user is equipped with $N$ antennas per RAP.
antennas. The baseband units (BBUs) are moved to a single BBU pool which are connected to the RAPs via high-speed fronthaul links, such that the BBU pool has access to the perfect channel state information (CSI) between the RAPs and the users, and the signals of all RAPs can be jointly processed. With a high density of geographically distributed RAPs, the access distances from each user to the RAPs varies significantly, and the distant RAPs have little contribution to improve the capacity. This motivates us to find a subset of RAPs that can provide near optimal sum rate performance.

In particular, let \( A \subseteq \mathcal{L} \) denotes the set of active RAPs, with \( |A| = A \). To utilize the multiplexing gains from the use of multiple user antennas, we assume that \( AN_c \geq N \). The received signal at user \( k \) can be then modeled as

\[
y_k = H_{k,A}x_k,A + \sum_{j \neq k} H_{k,A}x_{j,A} + z_k,
\]

where \( x_{k,A} \in \mathbb{C}^{AN_c \times 1} \) and \( y_k \in \mathbb{C}^{N \times 1} \) denote the transmit and receive signal vectors, respectively. \( H_{k,A} \in \mathbb{C}^{N \times AN_c} \) is the channel gain matrix between the active RAPs and user \( k \). \( z_k \) denotes the additive noise, which is modeled as a Gaussian random vector with zero mean and covariance \( \sigma^2 I \). With linear precoding, the transmit signal vector \( x_{k,A} \) can be expressed as

\[
x_{k,A} = T_{k,A}s_k,
\]

where \( T_{k,A} \in \mathbb{C}^{AN_c \times N} \) is the precoding matrix. \( s_k \in \mathbb{C}^{N \times 1} \) is the information bearing symbols. It is assumed that Gaussian codebook is used for each user at the transmitter, and therefore \( s_k \sim \mathcal{CN}(0, I) \). The transmit covariance matrix for user \( k \) can be then written as \( S_{k,A} = \mathbb{E}[x_{k,A}x_{k,A}^\dagger] \). It is easy to verify that \( S_{k,A} = T_{k,A}T_{k,A}^\dagger \). The sum rate can be written from (1) as

\[
R = \sum_k \log \det \left( I + \frac{H_{k,A}S_{k,A}H_{k,A}^\dagger}{\sigma^2 I + \sum_{j \neq k} H_{k,A}S_{j,A}H_{k,A}^\dagger} \right),
\]

Note that it is difficult to find the optimal linear precoder that maximizes the sum rate \( R \) due to the non-convexity of (3), and the mutual coupling of the transmit and receive beamformers makes it difficult to jointly optimize the beamformers [20]. In this paper, we assume that block diagonalization (BD) [15] is adopted, where an interference-free block channel is obtained by projecting the desired signal to the null space of the channel gain matrices of the other users, such that \( H_{k,A}x_{j,A} = 0 \), or equivalently \( H_{k,A}S_{j,A}H_{k,A}^\dagger = 0 \) for all \( j \neq k \). With BD, the sum rate can be obtained as

\[
R = \sum_k \log \det \left( I + \frac{1}{\sigma^2} H_{k,A}S_{k,A}H_{k,A}^\dagger \right).
\]

As the signals come from more than one RAPs, they need to satisfy a set of per-RAP power constraint, i.e.,

\[
\text{Tr} \left( B_l \sum_k S_{k,A} \right) \leq P_{l,\max}, \forall l \in A,
\]

where \( B_l \) is a positive definite matrix that represents the power constraint for the \( l \)-th RAP.
where $P_{l,\text{max}}$ denotes the per-RAP power constraint. $B_l = \text{diag}\{b_l\}$ is a diagonal matrix, whose diagonal entries is defined as

$$b_l = [0, \cdots, 0, 1, \cdots, 1, 0, \cdots, 0]. \quad (6)$$

This paper focuses on the tradeoff between the sum rate and the group sparsity. In particular, the problem can be formulated as the following optimization problem:

$$\begin{align*}
\text{maximize} & \quad \sum_k \log_2 \det \left( I + \frac{1}{\sigma^2} H_{k,A} S_k,A H_{k,A}^H \right) - \eta|A| \\
\text{s.t.} & \quad H_{j,A} S_{k,A} H_{j,A} = 0 \quad \forall j, k \in K, j \neq k \quad (7a) \\
& \quad \text{Tr} \left\{ B_l \sum_k S_{k,A} \right\} \leq P_{l,\text{max}} \quad \forall l \in A \quad (7b) \\
& \quad S_{k,A} \succeq 0 \quad \forall k \in K \quad (7c)
\end{align*}$$

where (7b) is the zero-forcing (ZF) constraint, which ensures that the inter-user interference can be completely eliminated at the optimum. $\eta \geq 0$ is the tradeoff constant, which controls the sparsity of the solution, and thus the number of active RAPs. With $\eta = 0$, the problem reduces to a BD precoder optimization problem with per-RAP power constraint. The group sparsity can be improved by assigning a larger $\eta$.

Note that the optimization problem (7) needs to jointly determine the subset $A$ and design the transmit covariance matrices $S_{k,A}$ for $K$ users, which is a combinatorial optimization problem and is NP-hard. A brute-force solution to a combinatorial optimization problem like (7) is by exhaustive search. Specifically, we must check all possible combinations of the active RAPs. For each combination, we must search for the optimal $\{S_{k,A}\}$ that satisfies the constraints (7b-7d). In the end, we pick out the combination that maximizes the sum rate. However, the complexity grows exponentially with $L$, which can not be applied to real-world application. Instead, we use the concept of $\ell_0$ norm to reformulate problem (7). In particular, define $\omega \in \mathbb{R}^{1 \times L}$ as

$$\omega = \left[ \text{Tr}\{B_1 \sum_k S_k\}, \text{Tr}\{B_2 \sum_k S_k\}, \cdots, \text{Tr}\{B_l \sum_k S_k\}, \cdots, \text{Tr}\{B_L \sum_k S_k\} \right]^T, \quad (8)$$

where $S_k = \mathbb{E}\left[ x_k x_k^H \right]$ with $x_k \in \mathbb{C}^{M \times 1}$ denoting the transmit signal vector from all RAPs in $L$ to user $k$. The $(N_c(l-1)+1)$-th to the $N_c l$-th entries of $x_k$ are zero if RAP $l$ is inactive. It is clear that the $l$-th entry $\omega_l = \text{Tr}\{B_l \sum_k S_k\}$ is the transmit power of RAP $l$, which is non-zero if and only if RAP $l \in A$. It is easy to verify that $\|\omega\|_0 = |A|$. We then have the following lemma:
Theorem 1 Problem (7) is equivalent to the following optimization problem:

\[
\begin{align*}
\text{maximize} & \quad \sum_k \log_2 \det \left( I + \frac{1}{\sigma^2} H_k S_k H_k^\dagger \right) - \eta \|\omega\|_0 \\
\text{s.t.} & \quad H_j S_k H_j = 0 \quad \forall j, k \in \mathcal{K}, j \neq k \\
& \quad \text{Tr} \left\{ B_l \sum_k S_k \right\} \leq P_{l,\text{max}} \quad \forall l \in \mathcal{L} \\
& \quad S_k \succeq 0 \quad \forall k \in \mathcal{K}
\end{align*}
\]  

(9)

where \( H_k \in \mathbb{C}^{N \times M} \) is the channel gain matrix from all RAPs in \( \mathcal{L} \) to user \( k \), \( k \in \mathcal{K} \).

Proof Please refer to Appendix A for detailed proof.

Lemma 1 indicates that instead of searching over the possible combinations of \( \mathcal{A} \) and then optimizing according to the corresponding channel gain matrices \( \{ H_k, A \} \) (7) can be solved based on the channel between the users and all RAP antennas, i.e., \( \{ H_k \} \). However, the problem (7) is non-convex due to the existence of the \( \ell_0 \) norm, making it difficult to find the global optimal solution.

In compressive sensing theory, the \( \ell_0 \) norm is usually replaced by a \( \ell_1 \) norm, and sparse solution can be achieved. However, simply substituting \( \|\omega\|_0 \) by \( \|\omega\|_1 \) in (7) will not necessarily produce sparse solution in general, as \( \|\omega\|_1 \) equals the sum power consumption instead of the number of non-zero entries. By replacing \( \|\omega\|_0 \) by \( \|\omega\|_1 \), the transmit power at all RAPs still tend to satisfy the power constraints with equality at the optimum, leading to a non-sparse solution. In this paper, we propose to solve (7) heuristically by iteratively relaxing the \( \ell_0 \) norm as a weighted \( \ell_1 \) norm. In particular, at the \( t \)-th iteration, the \( \ell_0 \) norm \( \|\omega^{(t)}\|_0 \) is approximated by

\[
\|\omega^{(t)}\|_0 \approx \sum_{l=1}^L \beta_l^{(t)} \|\omega^{(t)}\|_1,
\]

(10)

where \( \beta_l^{(t)} = \frac{1}{\text{Tr}(B_l \sum_k S_k^{(t-1)})} \), with \( 0 < \epsilon \ll \min(\{P_{l,\text{max}}\}_{l \in \mathcal{L}}) \) denoting a small positive constant. (7) can be then reformulated as

\[
\begin{align*}
\text{maximize} & \quad \sum_k \log_2 \det \left( I + \frac{1}{\sigma^2} H_k S_k H_k^\dagger \right) - \text{Tr} \left\{ \Psi^{(t)} \sum_k S_k \right\} \\
\text{s.t.} & \quad \text{Constraints in (9)},
\end{align*}
\]

(11)

where

\[
\Psi^{(t)} = \eta \sum_l \beta_l^{(t)} B_l.
\]

(12)

It is clear that problem (11) is convex, and can be solved by standard convex optimization techniques, e.g., interior point method [21], which, however, is typically slow. In fact, by utilizing the structure of BD precoding, the problem can be efficiently solved by its Lagrangian dual.
In this section, the algorithm to solve the group sparse linear precoding problem will be presented. To solve (11), it is desirable to remove the set of ZF constraints in the first place. It has been proved in [22] that the optimal solution for BD precoding with per-RAP constraint is given by

\[ S_k = \tilde{V}_k Q_k \tilde{V}_k^\dagger, \]  

(13)

where \( Q_k \succeq 0 \). \( \tilde{V}_k \) is given from the singular value decomposition (SVD) of \( G_k = [H_k^T, \ldots, H_{k-1}^T, H_k^T, \ldots, H_K^T]^T \) as

\[ G_k = U_k \Sigma_k [V_k, \tilde{V}_k]^\dagger, \]  

(14)

where \( \tilde{V}_k \in \mathbb{C}^{M \times (M - N(K - 1))} \) is the last \( M - N(K - 1) \) columns of the right singular matrix of \( G_k \). It is easy to verify that \( H_j S_k H_j^\dagger = 0 \) for all \( j \neq k \).

Therefore, by substituting (13) into (11), the problem reduces to

\[
\begin{align*}
\text{maximize} & \quad \sum_k \log_2 \det \left( I + \frac{1}{\sigma^2} H_k \tilde{V}_k Q_k \tilde{V}_k^\dagger H_k^\dagger \right) - \sum_{k \in K} \text{Tr} \left\{ \Psi(t) \tilde{V}_k Q_k \tilde{V}_k^\dagger \right\} \\
\text{s.t.} & \quad \sum_k \text{Tr} \left\{ B_l \tilde{V}_k Q_k \tilde{V}_k^\dagger \right\} \leq P_{l, \text{max}} \quad \forall \ l \in \mathcal{L} \\
& \quad Q_k \succeq 0 \quad \forall k \in \mathcal{K}
\end{align*}
\]  

(15)

Note that (15) is a convex problem, its Lagrangian dual can be written as

\[
L (\{Q_k\}, \lambda) = \sum_k \log_2 \det \left( I + \frac{1}{\sigma^2} H_k \tilde{V}_k Q_k \tilde{V}_k^\dagger H_k^\dagger \right) \\
- \sum_{k \in K} \text{Tr} \left\{ \Psi(t) \tilde{V}_k Q_k \tilde{V}_k^\dagger \right\} \\
- \sum_l \lambda_l \left( \sum_k \text{Tr} \left\{ B_l \tilde{V}_k Q_k \tilde{V}_k^\dagger \right\} - P_{l, \text{max}} \right)
\]  

(16)

where \( \lambda = [\lambda_l]_{l=1}^{L} \), with \( \lambda_l \geq 0 \) denoting the Lagrangian dual variables. The Lagrangian dual function can be given as

\[ g(\lambda) = \max_{Q_k \succeq 0} L (\{Q_k\}, \lambda). \]  

(17)

We can then obtain the Lagrangian dual problem of (15) as

\[ \min_{\lambda \geq 0} g(\lambda). \]  

(18)

Since the problem (15) is convex and satisfies the Slater’s condition, strong duality holds. The respective primal and dual objective values in (15) and (18) must be equal at the global optimum, and the complementary slackness must hold at the optimum [21], i.e.,

\[ \lambda_l^* \left( \sum_k \text{Tr} \left\{ B_l \tilde{V}_k Q_k \tilde{V}_k^\dagger \right\} - P_{l, \text{max}} \right) = 0, \ l = 1, \ldots, L, \]  

(19)
where \( \{Q_k^*\} \) and \( \{\lambda_t^*\} \) are the optimal primal and dual variables, respectively. As BD requires that the total number of transmit antennas should be equal or greater than the total number of user antennas, the minimum number of active RAPs should be \( A_{\min} = \left\lceil \frac{K}{N} \right\rceil \). We can then conclude from (14) that the maximum number of positive \( \lambda_t^* \)'s is

For fixed \( \lambda \), the Lagrangian dual function \( g(\lambda) \) can be obtained by solving

\[
\max_{Q_k \succeq 0} \sum_k \log_2 \det \left( I + \frac{1}{\sigma^2} H_k \tilde{V}_k Q_k \tilde{V}_k^\dagger \right) - \sum_k \text{Tr} \left\{ \Omega \tilde{V}_k Q_k \tilde{V}_k^\dagger \right\}, \quad (20)
\]

where \( \Omega = \Psi^{(t)} + \sum_l \lambda_l B_l \). By noting that \( \text{Tr} \{XY\} = \text{Tr} \{YX\} \), we have

\[
\text{Tr} \left\{ \Omega \tilde{V}_k Q_k \tilde{V}_k^\dagger \right\} = \text{Tr} \left\{ \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{1/2} Q_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{1/2} \right\}. \quad (21)
\]

Let \( \tilde{Q}_k = \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{1/2} Q_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{1/2} \). (20) can be written as

\[
\max_{Q_k \succeq 0} \sum_k \log_2 \det \left( I + \frac{1}{\sigma^2} H_k \tilde{V}_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{-1/2} \tilde{Q}_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{-1/2} \tilde{V}_k^\dagger H_k \right) - \sum_k \text{Tr} \left\{ \tilde{Q}_k \right\} \quad (22)
\]

It is clear that for given \( \lambda \), the problem (22) can be solved from \( K \) uncoupled subproblems:

\[
\max_{Q_k \succeq 0} \log_2 \det \left( I + \frac{1}{\sigma^2} H_k \tilde{V}_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{-1/2} \tilde{Q}_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{-1/2} \tilde{V}_k^\dagger H_k \right) - \sum_k \text{Tr} \left\{ \tilde{Q}_k \right\} \quad (23)
\]

Let us then solve the subproblems (23). By introducing the (reduced) SVD:

\[
H_k \tilde{V}_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{-1/2} = \hat{U}_k \Xi_k \tilde{V}_k^\dagger, \quad (24)
\]

where \( \hat{U}_k \in \mathbb{C}^{N \times N} \) and \( \tilde{V}_k \in \mathbb{C}^{N \times N} \) are unitary matrices, \( \Xi_k = \text{diag}(\xi_{k,1}, \cdots, \xi_{k,N}) \), with \( \xi_{k,n} \) denoting the \( n \)-th singular value of \( H_k \tilde{V}_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{-1/2} \). The optimal \( \tilde{Q}_k \) can be then obtained from the standard waterfiling algorithm [23]:

\[
\tilde{Q}_k^* = \tilde{V}_k^\dagger A_k \tilde{V}_k, \quad (25)
\]

where \( A_k = \text{diag}(\lambda_{k,1}, \cdots, \lambda_{k,N}) \), with

\[
\lambda_{k,n} = \left( 1 - \frac{\sigma^2}{\xi_{k,n}^2} \right)^+ \quad (26)
\]

where \( x^+ = \max(x, 0) \). The optimal \( S_k^* \) for given \( \lambda \) can be then obtained as

\[
S_k^*(\lambda) = \hat{V}_k \tilde{V}_k^\dagger \hat{U}_k \Xi_k \tilde{V}_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{-1/2} \tilde{V}_k^\dagger \Xi_k \tilde{V}_k \left( \tilde{V}_k^\dagger \Omega \tilde{V}_k \right)^{-1/2} \tilde{V}_k^\dagger \quad (27)
\]

With the optimal \( S_k^*(\lambda) \) is achieved for given \( \lambda \), we can then find the Lagrangian dual variables \( \lambda \) by the projected subgradient method. Projected subgradient methods following, e.g., the square summable but not summable step size rules,
Algorithm 1 Reweighted $\ell_1$ Norm Based Sparse Precoding Design

**Initialization:** Set iteration counter $t = 1$, Lagrangian dual variable $\lambda^{(0)} > 0$, $l = 1, \cdots, L$, and initial ellipsoid $P^{(0)} \in S^L$.

**Repeat:**

1. Calculate $\Psi^{(t)}$ and $S^{(t)}_k$, $k = 1, \cdots, K$, according to (12) and (27), respectively.
2. Update $\{\lambda_l\}$ according to (28).
3. Update the iteration counter $t := t + 1$ and $r = |\sum_l \lambda_l \sum_k \text{Tr}\{B_l \tilde{V}_k Q_k \tilde{V}_k^\dagger\}|$. 

**Stop** if $r < \varepsilon$, where $\varepsilon$ is a pre-defined tolerance threshold.

have been proved to converge to the optimal values [24]. In particular, a subgradient of $g(\lambda)$ with respect to $\lambda_l$ is $P_{l,\text{max}} - \sum_k \text{Tr}\{B_l \tilde{V}_k Q_k \tilde{V}_k^\dagger\}$. With a step size $\delta_t$, the dual variables can be updated as

$$\lambda_l^{(t+1)} = \max \left\{\lambda_l^{(t)} - \delta_t \left( P_{l,\text{max}} - \sum_k \text{Tr}\{B_l \tilde{V}_k Q_k \tilde{V}_k^\dagger\} \right), 0 \right\}, \quad \forall l = 1, \cdots, L.$$  

(28)

From the complementary slackness in (19), a stopping criterion for updating (28) can be

$$|\sum_l \lambda_l \sum_k \text{Tr}\{B_l \tilde{V}_k Q_k \tilde{V}_k^\dagger\}| < \varepsilon.$$  

(29)

Once the optimal $\lambda^*$ is obtained, the optimal precoding matrices $\{T_k^*\}$ can be achieved by using the fact that $S_k = T_k T_k^\dagger$ as

$$T_k^* = \tilde{V}_k \left( \tilde{V}_k^\dagger \Omega^* \tilde{V}_k \right)^{-1/2} \tilde{V}_k^\dagger \Lambda_k^{1/2}.$$  

(30)

The optimal set $A$ and the corresponding precoding matrices $\{T_{k,A}\}$ can be obtained from $\{T_k^*\}$. The algorithm is summarized as Algorithm 1.

4 Simulation Results

In this section, simulation results are presented to illustrate the results in this paper. We assume that the channel gain matrices $\{H_{k,l}\}$ from RAP $l$ to user $k$ are independent over $k$ and $l$ for all $l \in L$ and $k \in K$, and all entries of $H_{k,l}$ are independent and identically distributed (iid) complex Gaussian random variables with zero mean and variance $\gamma_{k,l}^2$. The path-loss model from RAP $l$ to the user $k$ is

$$\text{PL}_{k,l}(\text{dB}) = 128 + 37.6 \log_{10} D_{k,l},$$  

(31)

$\forall l \in L$ and $\forall k \in K$, where $D$ is in the unit of kilometer. The large-scale fading coefficient from RAP $l$ to user $k$ can be then obtained as

$$\gamma_{k,l}^2 = 10^{-\text{PL}_{k,l}(\text{dB})/10}.$$  

(32)
The transmit power constraint at each RAP are assumed to be identical, which is set to be $P_{l,\text{max}} = -40\text{dBm/Hz}$, $\forall l \in \mathcal{L}$, and the noise variance is set to be $\sigma^2 = -162\text{dBm/Hz}$.

We consider the case that $L = 10$ RAPs with $N_c = 2$ antennas each and $K = 2$ users with $N = 3$ antennas each. The positions of users and RAPs and the corresponding small-scale fading coefficients are randomly generated. Fig. 2 shows how the number of active RAPs varies with iterations, where the tradeoff constant $\eta$ is set to be 0, 0.1 and 0.5. The number of active RAPs corresponds to $|A|$. Fig. 3 shows that the number of active RAPs decreases with $\eta$. Specifically, a RAP is said to be active if its transmit power $P_l \geq 10^{-5}P_{l,\text{max}}$. 

---

**Fig. 2** Number of active RAPs under a randomly generated channel realization. $L = 10$, $N_c = 2$, $K = 2$, $N = 3$.

**Fig. 3** Transmit power of the RAPs under a randomly generated channel realization. $L = 10$, $N_c = 2$, $K = 2$, $N = 3$, $\eta = 0.5$. 
with $\eta = 0$, the problem reduces to a BD precoder design with per-RAP constraint, and all RAPs are active. With $\eta = 0.5$, the sparsest solution can be achieved, i.e., $|A| = A_{\text{min}} = \lceil \frac{KN}{N_c} \rceil = 3$. We should mention that the value of $\eta$ that corresponds to the sparsest solution varies with the system configuration.

As Fig. 2 shows, the first several iterations lead to biggest improvement. As iterations go on, there is no further improvement after the 15-th iteration. Compared to the full cooperation case, i.e., $\eta = 0$, as $L - A_{\text{min}} = 7$ RAPs are switched off, 70% of the power consumption can be saved with, however, limited rate performance loss, which will be illustrated later in this section. Fig. 3 plots the transmit power distribution over all 10 RAPs. From Fig. 3 we can see that as the iterations progress, RAP 1, 7 and 9 form a serving cluster and transmit with almost the maximum power, while all the other RAPs eventually drop their transmit power to zero after 15 iterations.

Fig. 4 plots how the average sum rate varies with the number of active RAPs. The average sum rate is obtained by averaging over 20 realizations of small-scale fading and 30 realizations of the positions of RAPs and users. The results of exhaustive search is also presented for comparison, which is obtained by searching over all possible combinations of active RAPs, and compute its achievable sum rate by the algorithm given in [22]. For the proposed algorithm, we simulate a series of different $\eta$’s to get different points along the curve. As we can see from Fig. 4 our proposed algorithm achieves almost the same average sum rate as the exhaustive search, which verifies the optimality of our proposed algorithm. Fig. 4 further plots the average sum rate with a fixed number of $|A|$ instead of $L$ uniformly distributed antennas are deployed for comparison, which is denoted as “Fixed $L$” in Fig. 4. We can clearly see that the proposed algorithm can achieve much better rate performance over that with the same amount of transmit RAP. With 6 selected antennas, for instance, the group sparse precoder reduces the average sum rate for only 3 bit/s/Hz, whereas if only 6 antennas were installed instead of $L = 10$, an additional of 9 bit/s/Hz rate loss can be observed compared to the proposed algorithm. Moreover, the rate gap between the proposed algorithm and that with a fixed number of $L = |A|$ full cooperative RAPs further increases as the number of selected RAPs $|A|$ decreases. It highlights the importance of group sparse precoder design in C-RAN with a large number of distributed RAPs.

5 Conclusion

In this paper, we study the group sparse precoder design that maximizes the sum rate in a C-RAN under per-RAP power constraint. We show that the joint antenna selection and precoder design problem can be formulated into an $\ell_0$ norm problem, which is, however, combinatorial and NP-hard. Inspired by the theory of compressive sensing, we propose an approach that solves the problem via reweighted $\ell_1$ norm. Simulation results verify the optimality of our proposed algorithm in that it achieves almost the same performance as that obtained form the exhaustive search. Compared to full cooperation, the group sparse precoding can achieve a significant proportion of the maximum sum rate that achieved from full cooperation with, however, much fewer active RAPs, which highlights the importance to employ sparse precoding in C-RAN with ultra-dense RAPs.
Appendix

A Proof of Lemma 1

Define the channel gain matrix and the precoding matrix between user $k$ and the antennas of RAP $l$ as $H_{k,l}$ and $T_{k,l}$, respectively, $\forall k \in \mathcal{K}$ and $\forall l \in \mathcal{L}$. Let us reorder the channel gain matrices $H_k$ and the precoding matrices $T_k$ as

$$\tilde{H}_k = \begin{bmatrix} [H_{k,l}]_{l \in \mathcal{A}}, & [H_{k,l}]_{l \notin \mathcal{A}} \end{bmatrix},$$

and

$$\tilde{T}_k = \begin{bmatrix} [T_{k,l}^T]_{l \in \mathcal{A}}, & [T_{k,l}^T]_{l \notin \mathcal{A}} \end{bmatrix}^T,$$

respectively. We have

$$H_j S_k H_j = \tilde{H}_k \tilde{T}_k \tilde{T}_k^\dagger \tilde{H}_k^\dagger \overset{(a)}{=} \left( \sum_{l \in \mathcal{A}} H_{k,l} T_{k,l} \right) \left( \sum_{l \in \mathcal{A}} H_{k,l} T_{k,l}^\dagger \right)^\dagger$$

$$= H_{\mathcal{J},\mathcal{A}} S_{\mathcal{K},\mathcal{A}} H_{\mathcal{J},\mathcal{A}}^\dagger,$$

for all $j, k \in \mathcal{K}$, where (a) follows from the fact that $T_{k,l} = 0$ if RAP $l$ is not active, i.e., $l \notin \mathcal{A}$.

Let us then prove the positive semidefinite constraint. The reordered covariance matrix $S_k$ can be written as

$$S_k = \tilde{T}_k \tilde{T}_k^\dagger \overset{(b)}{=} \begin{bmatrix} T_{k,\mathcal{A}} T_{k,\mathcal{A}}^\dagger & T_{k,\mathcal{A}} [T_{k,l}]_{l \notin \mathcal{A}}^\dagger \\ [T_{k,l}]_{l \notin \mathcal{A}} T_{k,\mathcal{A}}^\dagger & [T_{k,l}]_{l \notin \mathcal{A}} [T_{k,l}]_{l \notin \mathcal{A}}^\dagger \end{bmatrix}$$

$$\overset{(b)}{=} \begin{bmatrix} T_{k,\mathcal{A}} T_{k,\mathcal{A}}^\dagger & 0 \\ 0 & 0 \end{bmatrix},$$

where (b) holds as $T_{k,l} = 0$ for all $l \notin \mathcal{A}$. By noting that $T_{k,\mathcal{A}} T_{k,\mathcal{A}}^\dagger = S_{k,\mathcal{A}} \succeq 0$, we immediately have $S_k \succeq 0$. (9) can be then proved by combining (35-36) and (7).
References

1. C-RAN: The Road Towards Green RAN. White paper, ver. 2.5.
2. Shi, Y., Zhang, J., Letaief, K.B., Bai, B., Chen, W. (2015): Large-scale convex optimization for ultra-dense cloud-ran. *IEEE Wireless Commun.*, 22(3), 84–91.
3. Liu, C., Zhang, L., Zhu, M., Wang, J., Cheng, L., Chang, G.K. (2013): A novel multi-service small-cell cloud radio access network for mobile backhaul and computing based on radio-over-fiber technologies. *J. Lightwave Technol.*, 31(17), 2869–2875.
4. Liu, Z., Dai, L. (2014): A comparative study of downlink mimo cellular networks with co-located and distributed base-station antennas. *IEEE Trans. Wireless Commun.*, 13(11), 6259–6274.
5. Liu, Z.: On the scaling behavior of the average rate performance of large-scale distributed MIMO systems. *IEEE Trans. Veh. Technol.* To appear.
6. Wang, J., Dai, L. (2015): Asymptotic rate analysis of downlink multi-user systems with co-located and distributed antennas. *IEEE Trans. Wireless Commun.*, 14(6), 3046–3058.
7. Zhao, J., Quek, T.Q.S., Lei, Z. (2013): Coordinated multipoint transmission with limited backhaul data transfer. *IEEE Trans. Wireless Commun.*, 12(6), 2762–2775.
8. Mehanna, O., Sidiropoulos, N.D., Giannakis, G.B. (2013): Joint multicast beamforming and antenna selection. *IEEE Trans. Signal Process.*, 61(10), 2660–2674.
9. (2013): Sparse beamforming for limited-backhaul network mimo system via reweighted power minimization. In: Proc. IEEE Globecom.
10. Zhang, Z., Xu, Y., Yang, J., Li, X., Zhang, D. (2015): A survey of sparse representation: Algorithms and applications. *IEEE Access*, 3, 490–530.
11. Dai, B., Yu, W. (2014): Sparse beamforming and user-centric clustering for downlink cloud radio access network. *IEEE Access*, 2, 1326–1339.
12. Shi, Y., Zhang, J., Letaief, K.B. (2014): Group sparse beamforming for green cloud-ran. *IEEE Trans. Wireless Commun.*, 13(5), 2809–2823.
13. Shi, Y., Zhang, J., Letaief, K.B. (2015): Robust group sparse beamforming for multicast green cloud-ran with imperfect csi. *IEEE Trans. Signal Process.*, 63(17), 4647–4659.
14. Telatar, E. (1999): Capacity of multi-antenna Gaussian channels. *Euro. Trans. Telecommun.*, 10(6), 585–595.
15. Spencer, Q., Swindlehurst, A., Haardt, M. (2004): Zero-forcing methods for downlink spatial multiplexing for multiuser MIMO channels. *IEEE Trans. Signal Process.*, 52(2), 461–471.
16. Shen, Z., Chen, R., Andrews, J., Heath, R., Evans, B. (2006): Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization. *IEEE Trans. Signal Process.*, 54(9), 3658–3663.
17. Shen, Z., Chen, R., Andrews, J., Heath, R., Evans, B. (2007): Sum capacity of multiuser MIMO broadcast channels with block diagonalization. *IEEE Trans. Wireless Commun.*, 6(6), 2040–2045.
18. Shin, S., Kwak, J.S., Heath, R., Andrews, J. (2008): Block diagonalization for multi-user MIMO with other-cell interference. *IEEE Trans. Wireless Commun.*, 7(7), 2671–2681.
19. Ravindran, N., Jindal, N. (2008): Limited feedback-based block diagonalization for the MIMO broadcast channel. *IEEE J. Sel. Areas Commun.*, 26(8), 1473–1482.
20. Cai, D.W.H., Quek, T.Q.S., Tan, C.W. (2011): A unified analysis of max-min weighted sinr for MIMO downlink system. *IEEE Trans. Signal Process.*, 59(8), 3850–3862.
21. Boyd, S., Vandenberghe, L. (2004): *Convex Optimization*. Cambridge University Press.
22. Zhang, R. (2010): Cooperative multi-cell block diagonalization with per-base-station power constraints. *IEEE J. Sel. Areas Commun.*, 28(9), 1435–1445.
23. Cover, T.M., Thomas, J.A. (2006): *Elements of Information Theory*. Wiley-Interscience.
24. Bertsekas, D.P. (2003): *Convex Analysis and Optimization*. Athena Scientific.
