

Energy-Efficient Joint Association and Precoding in Ultra-Dense C-RAN

Alireza Zamani*, Omid Taghizadeh[†] and Anke Schmeink*

*ISEK Research Group, RWTH Aachen University, Aachen, Germany

Email: {firstname.lastname}@isek.rwth-aachen.de

[†] Network Information Theory Group, Technische Universität Berlin, Germany

Email: {taghizadehmotlagh}@tu-berlin.de

Abstract—In this work, we consider the joint optimization of virtual cell association and the precoder design for a multi-user downlink cloud radio access network (C-RAN). The optimal design of cell associations and precoders are particularly relevant due to the anticipated ultra-dense networks suffering from heavy interference. As the resulting optimization is a non-convex and combinatorial problem, we propose an iterative low complexity algorithm, employing the recently developed penalty dual decomposition (PDD) framework. In contrast to predominantly used sparsity inducing methods, the proposed method incorporates the penalty/re-weighting strategy into the algorithm, which is then solved as a sequence of convex second-order cone programs. The aforementioned feature leads to a high convergence rate and subsequently a low computational cost for the proposed algorithm. Furthermore, the proposed algorithm meets the requirements to converge to a KKT solution. Numerical simulations verify the superior performance of our proposed method in ultra-dense networks, in comparison to commonly used methods, such as the reweighted ℓ_1 -norm, in terms of system energy efficiency.

Index Terms—Signal processing, convex optimization, mixed integer programming, cloud radio access network.

I. INTRODUCTION

The remote radio heads (RRHs) in a C-RAN are connected via fast fronthaul links to the central unit (CU) which contains a baseband unit (BBU) pool. The availability of global CSI and high computational power at the CU allows RRHs to form a cooperative cluster. This paradigm replaces the traditionally strict concept of cell association by one where the CU groups the antennas to serve a user by selecting multiple RRHs. For this purpose, the term virtual cell association is more appropriate for referring to the RRHs that are selected to serve a given user. This work focuses on minimizing the power consumption of a network, including the power required in the fronthaul, via joint optimization of the virtual association and precoding. This joint optimization problem is motivated by the following arguments: I) cooperative techniques incur additional fronthaul load and power which must be accounted for II) conventional association schemes, e.g., based on distance or SNR do not suit the ultra-dense networks as they cause load imbalance III) fluctuating SINR demands or channel conditions may justify optimizing the user association on a shorter time-scale than traditionally.

Motivated by the advancements in compressed sensing, the majority of existing works use ℓ_0 -norm approximation/relaxation to describe the relationship between the precoding vectors and the cell association. The authors in [1]–[7] all proposed variations of sparsity inducing norms, e.g., reweighted ℓ_1 -norm, for joint precoding and cell association. This is approached by formulating a power minimization or rate maximization problem, where the association is expressed via the ℓ_1 -norm of the precoding vectors and then solving the problem as cone programming or via semi-definite relaxation. Other works [8], [9], proposed frameworks based on mixed-integer programming for joint user association and beamforming in a C-RAN. The authors in [10]–[12] proposed sparse approximations/relaxations to minimize the power consumption and provide robustness. In our previous work [13], a mixed integer-second order cone programming (MI-SoCP) framework was proposed for a C-RAN with limited CSI and fronthaul capacity. While MIPs offer a high-quality solution, they are computationally inefficient to solve. On the other hand, iterative sparsity inducing norm methods benefit from low computation complexity while offering sub-optimal solutions. Furthermore, as sparsity inducing norms do not explicitly include the association, they require careful pruning and post-processing which is not only a cumbersome task but can also degrade the quality of the solution.

We propose a novel algorithm for joint virtual cell association and precoding optimization. This is the first work to consider the application of PDD to this class of combinatorial problems. The proposed method replaces the NP-hard MIP problem by a sequence of convex problems, thus offering a significant reduction in computational complexity. Furthermore, the proposed algorithm has the significant advantage of reaching a KKT solution of the substitute problem to the original MIP. Simulation results indicate that the proposed method outperforms the reweighted ℓ_1 -norm, in ultra-dense networks.

A. Notation:

Uppercase calligraphic, lowercase and uppercase bold letters denote sets, vectors and matrices, respectively. \mathbb{C} , \Re and \Im ,

denote the set of complex numbers, the real and imaginary parts of a complex number, respectively. The Hermitian transpose and Euclidean norm are denoted by $(\cdot)^H$, $\|\cdot\|_2$, respectively. We use $\text{vec}\{A\}$ to show the vector concatenation of the columns of A .

II. SYSTEM MODEL & PROBLEM FORMULATION

In this section, we describe the system model of the C-RAN and formulate the corresponding optimization problem.

We consider an ultra-dense C-RAN, populated by I RRHs and J co-channel downlink users. The RRHs, which can encode and decode, are assumed to perform channel estimation and transfer the information to the CU which then returns the designed precoding vectors and associations. The set of RRHs and users is described by $\mathcal{I} = \{1, \dots, I\}$ and $\mathcal{J} = \{1, \dots, J\}$, respectively. Each RRH i is equipped with M antennas and has a maximum transmit power of P_i^{\max} . Users are assumed to be equipped with a single antenna and have a desired QoS given as the minimum required average SINR and rate, denoted by γ_j and r_j , $1 \leq j \leq J$. The channel between the i -th RRH and the j -th user is represented by $\mathbf{h}_{ij} \in \mathbb{C}^{M \times 1}$ and is modeled as uncorrelated block flat fading with Rayleigh distribution. The precoding vector between the i -th RRH and the j -th user is denoted by $\mathbf{w}_{ij} \in \mathbb{C}^{M \times 1}$. For ease of notation, we define $\mathbf{h}_j \in \mathbb{C}^{M_T \times 1}$ and $\mathbf{w}_j \in \mathbb{C}^{M_T \times 1}$ as the global channel and precoding vectors of the j -th user, where $M_T = I \cdot M$. The data symbol intended for the j -th user is represented as $s_j \sim \mathcal{CN}(0, 1)$. The complex additive white Gaussian noise (AWGN) is denoted by $z_j \sim \mathcal{CN}(0, \sigma_j^2)$. For the described system, the received signal by the j -th user can be written as

$$y_j = \mathbf{h}_j^H \mathbf{w}_j s_j + \sum_{q \neq j} \mathbf{h}_j^H \mathbf{w}_q s_q + z_j, \quad j, q \in \mathcal{J}. \quad (1)$$

The virtual cell association between the i -th RRH and the j -th user is given by $\alpha_{ij} \in \{0, 1\}$, where a one indicates an active association and vice versa. Note that an active association not only has the implication that the corresponding precoder is non-zero, but also influences fronthaul load.

A. Network Power Consumption Model

The power consumption considered in the network is comprised of the transmit power at each RRH and the power consumption of the fronthaul links. As the signaling overhead takes place on a larger time scale, we assume the fronthaul load is dominated by the users' data.

The fronthaul model used follows that of [14] for microwave links with the fronthaul power consumption given as

$$P_i^{FH} = \frac{\sum_j \alpha_{ij} r_j}{C_i} P_{i,max}^{FH}, \quad i \in \mathcal{I}, \quad (2)$$

where C_i and $P_{i,max}^{FH}$ denote the capacity of the fronthaul communication channel and the power dissipation, respectively.

The total network power consumption is expressed as

$$P_{Tot}(\{\alpha_{ij}\}, \{\mathbf{w}_j\}, \{P_i^{FH}\}) = \sum_{i,j} \|\mathbf{w}_{ij}\|_2^2 + \sum_i P_i^{FH}, \quad (3)$$

$$i \in \mathcal{I}, j \in \mathcal{J}.$$

B. Problem Formulation

We begin by formulating the joint virtual cell association and precoder optimization as a MI-SoCP problem as shown in [13]

P1 :

$$\begin{aligned} \min_{\{\alpha_{ij}\}, \{\mathbf{w}_j\}, \{P_i^{FH}\}} & P_{Tot}(\{\alpha_{ij}\}, \{\mathbf{w}_j\}, \{P_i^{FH}\}) \\ \text{s.t.} & (2), \\ & \|\mathbf{w}_{ij}\|_2^2 \leq \alpha_{ij} P_i^{\max}, \quad i \in \mathcal{I}, j \in \mathcal{J}, \quad (4a) \\ & \sum_j \|\mathbf{w}_{ij}\|_2^2 \leq P_i^{\max}, \quad i \in \mathcal{I}, \quad (4b) \\ & \left\| \begin{pmatrix} \text{vec}\{\mathbf{h}_j^H \mathbf{W}\} \\ \sigma_j \end{pmatrix} \right\|_2 \leq \sqrt{1 + \frac{1}{\gamma_j} \mathbf{h}_j^H \mathbf{w}_j}, \\ & \quad j \in \mathcal{J}, \quad (4c) \\ & \Im\{\mathbf{h}_j^H \mathbf{w}_j\} = 0, \quad j \in \mathcal{J}, \quad (4d) \\ & \alpha_{ij} \in \{0, 1\}, \quad i \in \mathcal{I}, j \in \mathcal{J}, \quad (4e) \end{aligned}$$

where $\mathbf{W} \in \mathbb{C}^{M_T \times J}$ is the concatenation of the global precoding vector of all users. Note that (4a) relates the association and the precoding vectors while constraint (4b) represents the maximum transmit power constraint per RRH. Note that P1 is NP-hard and computationally expensive to solve.

III. PDD-BASED JOINT VIRTUAL CELL ASSOCIATION AND PRECODER OPTIMIZATION

In order to facilitate the implementation of the proposed PAP algorithm, we first propose a reformulation of P1.

A. Problem Transformation

As the complexity in solving P1 arises from the binary association constraint (4e), we first consider the following equivalent reformulation

$$\alpha_{ij} \in \{0, 1\} \equiv \alpha_{ij}^2 - \alpha_{ij} = 0. \quad (5)$$

Although (5) is still non-convex, we aim to obtain a tractable structure for the PDD algorithm by using an auxiliary variable β_{ij} , and expressing the above constraint as

$$\alpha_{ij} \beta_{ij} = 0, \quad (6a)$$

$$\alpha_{ij} + \beta_{ij} = 1. \quad (6b)$$

This facilitates a reformulation of the optimization problem as

P2 :

$$\begin{aligned} \min_{\{\alpha_{ij}\}, \{\beta_{ij}\}, \{\mathbf{w}_j\}, \{P_i^{FH}\}} & P_{Tot}(\{\alpha_{ij}\}, \{\mathbf{w}_j\}, \{P_i^{FH}\}) \\ \text{s.t.} & (2), (4a), (4b), (4c), (4d), \\ & (6a), (6b), \quad i \in \mathcal{I}, j \in \mathcal{J}. \end{aligned}$$

We note that in P2, the constraints related to the association are still expressed using α_{ij} . While P2 is not jointly convex over all the optimization variables, due to constraint (6a), the optimization variables can be separated and updated in a block-wise manner in adherence with the PDD algorithm.

B. PDD Method

In this subsection, we briefly explain the general PDD method [15], which is used in our proposed framework. For this purpose, we consider the following optimization problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{z}(\mathbf{x}) = \mathbf{0}, \\ & \mathbf{g}_i(\mathbf{x}_i) \leq \mathbf{0}, \quad \forall i \in \{1, \dots, n\}, \end{aligned} \quad (8)$$

where $f(\mathbf{x}) \in \mathcal{R}$, is a scalar and continuously differentiable function, while $\mathbf{z}(\mathbf{x}) \in \mathcal{R}^m$ is a vector of m continuously differentiable functions, but possibly coupled and non-convex, and $\mathbf{g}_i(\mathbf{x}_i) \in \mathcal{R}^n$ is a vector of n differentiable functions. Note that if the coupled constraint $\mathbf{z}(\mathbf{x}) = \mathbf{0}$ did not exist, then a KKT solution to the above problem could be obtained by employing block coordinate descent methods such as BSUM. The PDD method aims to resolve the coupled equality constraint via a double loop algorithm, where outer loop updates the dual variables and the penalty weights and the inner loop solves the Augmented Lagrangian (AL) optimization problem corresponding to (8) as shown below

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) + \boldsymbol{\eta}^T \mathbf{z}(\mathbf{x}) + \frac{1}{2\rho} \|\mathbf{z}(\mathbf{x})\|^2 \\ & = f(\mathbf{x}) + \frac{1}{2\rho} \|\mathbf{z}(\mathbf{x}) + \rho \boldsymbol{\eta}\|^2, \\ \text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i) \leq \mathbf{0}, \quad \forall i \in \{1, \dots, n\}, \end{aligned} \quad (9)$$

where ρ and $\boldsymbol{\eta}$ denote the penalty weight and the dual variable, respectively. It is clear that for a sufficiently large $\frac{1}{2\rho}$, the solution of (9) is identical to (8). However, achieving a significantly large weight with standard penalty methods may lead to slow convergence, or ill-conditioned problems. For this purpose, when the violations are larger than a threshold $\mathbf{z}(\mathbf{x}) > \epsilon$, PDD updates the penalty weight as $\rho^{k+1} = c\rho^k$. On the other hand, when the violations are small, $\mathbf{z}(\mathbf{x}) \leq \epsilon$, the dual variables are updated using dual ascent [16] as $\boldsymbol{\eta}^{k+1} = \boldsymbol{\eta}^k + \frac{1}{\rho^k} \mathbf{z}(\mathbf{x})$.

C. Proposed PDD-based Association and Precoding (PAP) Algorithm

The corresponding AL problem for P2 is derived as follows

P3 :

$$\begin{aligned} \min_{\{\alpha_{ij}\}, \{\beta_{ij}\}, \{\mathbf{w}_j\}, \{P_i^{FH}\}} \quad & P_{Tot}(\{\alpha_{ij}\}, \{\mathbf{w}_j\}, \{P_i^{FH}\}) + \frac{1}{2\rho^k} \sum_{i,j} \phi_{ij}^k \\ \text{s.t.} \quad & (2), (4a), (4b), (4c), (4d), \quad i \in \mathcal{I}, j \in \mathcal{J}, \end{aligned} \quad (10)$$

where k denotes the k -th outer iteration and $\phi_{ij}^k = (|\alpha_{ij}\beta_{ij} + \rho^k \lambda_{ij}^k|^2 + |\alpha_{ij} + \beta_{ij} - 1 + \rho^k \mu_{ij}^k|^2)$. Furthermore, ρ^k , λ_{ij}^k and μ_{ij}^k denote the penalty weight and the Lagrangian multipliers associated with constraints (6a) and (6b), respectively. We can solve the AL in P3 by separating the variables into two blocks, hence, creating two sub-problems: 1) while fixing $\{\beta_{ij}\}$ solve P3 for $\{\alpha_{ij}\}$, $\{\mathbf{w}_j\}$ and $\{P_i^{FH}\}$. 2) fixing $\{\alpha_{ij}\}$, $\{\mathbf{w}_j\}$ and $\{P_i^{FH}\}$ solve P3 for $\{\beta_{ij}\}$. It is worth

mentioning that optimizing the AL over each variable block is convex, furthermore, the solution of the second sub-problem can be obtained in closed form as

$$\beta_{ij} = \frac{1 - \alpha_{ij}}{\alpha_{ij}^2 + 1}. \quad (11)$$

The proposed PAP framework is summarized in Algorithm 1.

D. Initialization

For the first iteration of the algorithm, it is necessary to initialize $\beta_{ij} \in [0, 1]$. However, since the algorithm is capable of optimizing the associations $\{\alpha_{ij}\}$, a heuristic initialization is sufficient. Nevertheless, it is obvious that a good initial point can lead to faster convergence and/or a better solution. For this purpose, given the channel vectors \mathbf{h}_j , we set all associations to one and use zero-forcing precoding together with water-filling power allocation [17] to satisfy the QoS constraints γ_j . From this we then calculate the effective SNR between each RRH and user pair and initialize β_{ij} to favor the strongest links as follows

$$\beta_{ij} = 1 - \frac{\text{SNR}_{ij}}{\max_i \text{SNR}_{ij}}, \quad (12)$$

where SNR_{ij} denotes the SNR between the i -th RRH and j -th user.

Algorithm 1: Proposed PAP algorithm

Initialization: $\lambda_{ij}^k = 0$, $\mu_{ij}^k = 0$, $\epsilon = 0.1$, $\rho^0 = 1e^{-3}$
 $c = 0.25$, $k = 1$, β_{ij} see Section III-D

repeat

repeat

 Solve P3 over $\{\alpha_{ij}\}$, $\{\mathbf{w}_j\}$ and $\{P_i^{FH}\}$ via off-the-shelf solvers e.g., Gurobi [18].

 Compute β_{ij} via (11);

until Convergence of AL objective;

if $\max_{i,j} \{|\alpha_{ij}\beta_{ij}|, |\alpha_{ij} + \beta_{ij} - 1|\} \leq \epsilon$; **then**

$\lambda_{ij}^{k+1} = \lambda_{ij}^k + \frac{1}{\rho^k} (\alpha_{ij}\beta_{ij})$;

$\mu_{ij}^{k+1} = \mu_{ij}^k + \frac{1}{\rho^k} (\alpha_{ij} + \beta_{ij} - 1)$;

else

$\lambda_{ij}^{k+1} = \lambda_{ij}^k$;

$\mu_{ij}^{k+1} = \mu_{ij}^k$;

$\rho^{k+1} = c\rho^k$;

end

$k = k + 1$;

until Convergence of constraint violation;

E. Convergence & Complexity Analysis

P3 is a special case of the general problem studied in [15], where the PAP leads to a solution satisfying the KKT conditions, given that the inner loop converges to a stationary point of the AL. At each inner iteration, the AL problem is convex with respect to the block variables and can be solved optimally. In this regard, we achieve a monotonic decrease for

both the overall PDD objective for every inner iteration as well as the constraint violation for every outer iteration, as shown in Section IV via numerical evaluation. Furthermore, as the objective and constraints are smooth and decoupled, we meet the requirements specified in [19] for converging to a KKT solution. As the proposed framework achieves a KKT solution for P2, the solution is also inherently feasible for P1, as the obtained associations will satisfy the binary constraint.

The arithmetic complexity of the original problem presented in P1 is NP-hard, which makes it impractical to solve for large sizes. In contrast, since the update to $\{\beta_{ij}\}$ can be obtained in closed form, the complexity in our proposed algorithm comes from solving instances of P3 as second order cone programming (SoCP). Supposing the algorithm solves K instances of P3 to satisfy the convergence criteria the complexity is approximately $\mathcal{O}((IJM)^{3.5}K)$.

In terms of computational run-time, while P1 can be cast as a MI-SoCP, for which there are off-the-shelf solvers such as Gurobi, the run-time grows exponentially with the problem size. Therefore, comparing the computation run-time performance is not only dependent on the solver, but also impractical for larger problems where our proposed algorithm offers significant gains. Nevertheless, Table I. in Section IV offers an insight into the average run time of solving an instance of the SoCP in the PAP algorithm compared to the MI-SoCP.

IV. SIMULATION RESULTS & DISCUSSION

In this section, we present numerical simulations to evaluate the performance of the system under 3GPP LTE specifications. 3 RRHs and 4 downlink users are assumed to populate a C-RAN with a radius of 250 m. Each RRH is equipped with two antennas and has a maximum transmit power $P_i^{max} = 43$ dBm. The fronthaul link parameters are set as $C_i = 10$ Gbps and $P_{i,max}^{FH} = 50$ W. We assume a carrier frequency of 2 GHz and a channel bandwidth of 10 MHz in the access network. The channels follow an uncorrelated Rayleigh distribution with shadowing and $\sigma_j = -164$ dBm/Hz. Due to the limitation of space, we refer the reader to [13] for a detailed overview of the simulation parameters and setup. The simulations results are averaged over 200 channel realizations.

Table I. offers insight into the computation run time performance of the PAP algorithm.

TABLE I: Run time vs. complexity

Complexity	Average Run Time [s]		
	SoCP instance	Total PAP Convergence	MI-SoCP
#RRH = 2, #U _{ser} = 3	0.08	7.09	1.41
#RRH = 3, #U _{ser} = 4	0.12	10.26	54.53
#RRH = 4, #U _{ser} = 5	0.14	12.97	366.85
#RRH = 4, #U _{ser} = 7	0.18	16.91	–
#RRH = 5, #U _{ser} = 9	0.24	23.30	–

Figure 1 illustrates the reduction of the objective value of the AL in the PAP algorithm, with respect to the inner iterations, where convergence is typically achieved within 5 iterations. The convergence of sum violation corresponding to binary constraints (6b) and (6b) is shown in Fig. 2. We note that once

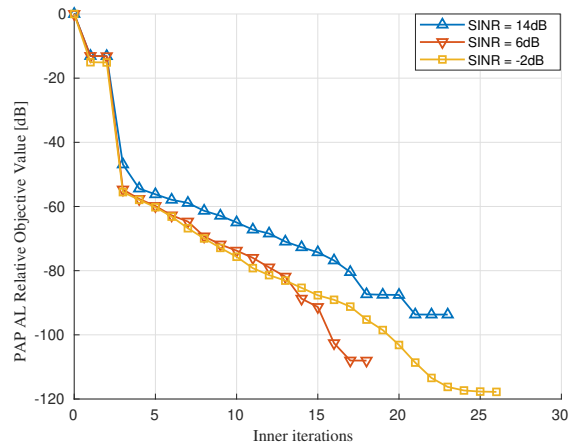


Fig. 1: Inner loop convergence behaviour of Algorithm 1.

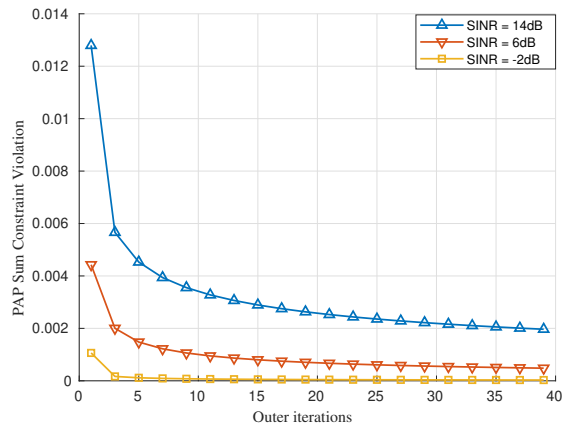


Fig. 2: Outer loop convergence behaviour of Algorithm 1.

the constraint violations are small enough, the algorithm can be terminated and a pruning step can be applied to obtain truly binary association variables. The fast convergence rate in both inner and outer loops demonstrate the low computation time required, which is due to the dynamic penalty update rules in the algorithm.

The total network power consumption is evaluated in Fig. 3. It is observed that the proposed algorithm offers significant power reduction compared to reweighted ℓ_1 -norm. More specifically, it achieves an improvement of over 50% at low data rate requirements and almost 10% for higher data rates and benefits from a smaller optimality gap.

The fronthaul load imposed on the network is studied in Fig. 4, where it can be deduced that the PAP algorithm consumes less capacity than reweighted ℓ_1 -norm. The inferior performance of reweighted ℓ_1 -norm is explained by the fact that in ultra-dense setups, the majority of users may receive a reasonable effective SINR from multiple RRHs as opposed to a single dominant node. Consequently, achieving sparsity by a fixed re-weighting rule is more difficult.

The superior performance of the PAP algorithm, relative

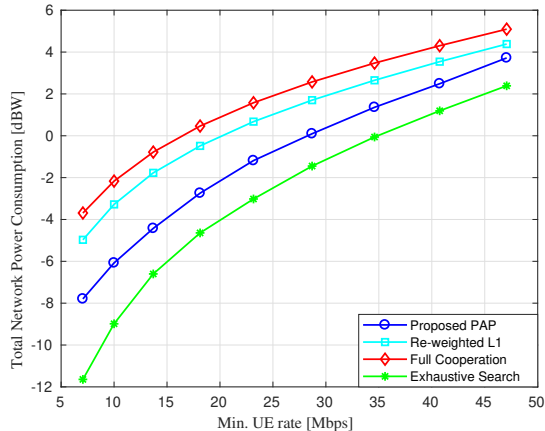


Fig. 3: Power consumption comparison of different methods.

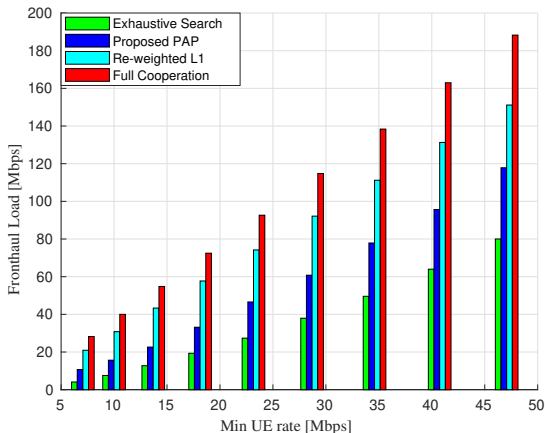


Fig. 4: Fronthaul load comparison of different methods.

to reweighted ℓ_1 -norm, is explained by its ability to more accurately model the relationship between the association and precoder as defined in constraint (4a). This is realized by the inclusion of the binary penalty terms in the AL, that promote the association to binary values, and the dynamic update mechanism incorporated in the algorithm, that depends on the constraint violation. In contrast, reweighted ℓ_1 -norm has a fixed update rule and does not directly include the association variables. Furthermore, it typically relies on more advanced pruning as the obtained solutions are not as close to binary values, especially in ultra-dense networks where a user can receive a high effective SINR from multiple RRHs.

V. CONCLUSION

In this work, we proposed a new energy-efficient algorithm for the joint optimization of virtual cell associations and precoders. The aforementioned problem is known to be non-convex and combinatorial and is widely solved via sparsity inducing norms. Our proposed framework employs the recently developed PDD method to solve a sequence of convex problems. The

proposed framework dynamically adjusts the penalty weights and dual variables to promote binary values, using an update mechanism incorporated in the algorithm. The PAP algorithm displays a fast convergence to a KKT solution, and a similar computation complexity to reweighted ℓ_1 -norm, while achieving a significant improvement in terms of power consumption.

REFERENCES

- [1] F. Zhang, Q. Wu, H. Wang, and Y. Shi, " ℓ_2 -Box Optimization for Green Cloud-RAN via Network Adaptation," 2017.
- [2] B. Dai and W. Yu, "Energy Efficiency of Downlink Transmission Strategies for Cloud Radio Access Networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 1037–1050, 2016.
- [3] —, "Sparse Beamforming and User-Centric Clustering for Downlink Cloud Radio Access Network," *IEEE Access*, vol. 2, pp. 1326–1339, 2014.
- [4] V. N. Ha, L. B. Le, and N. Ðào, "Coordinated Multipoint Transmission Design for Cloud-RANs with Limited Fronthaul Capacity Constraints," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7432–7447, 2016.
- [5] Y. Shi, J. Cheng, J. Zhang, B. Bai, W. Chen, and K. B. Letaief, "Smoothed ℓ_p -Minimization for Green Cloud-RAN with User Admission Control," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 1022–1036, 2016.
- [6] Y. Shi, J. Zhang, W. Chen, and K. B. Letaief, "Enhanced Group Sparse Beamforming for Green Cloud-RAN: A Random Matrix Approach," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, April 2018.
- [7] J. Zhao, T. Q. S. Quek, and Z. Lei, "Coordinated Multipoint Transmission with Limited Backhaul Data Transfer," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2762–2775, 2013.
- [8] A. Zamani, O. Taghizadeh, and A. Schmeink, "A Novel Joint Precoding and Association Optimization Framework for C-RANs with Restricted Fronthauls," in *WSA; 22nd International ITG Workshop on Smart Antennas*, March 2018.
- [9] Y. Cheng, M. Pesavento, and A. Philipp, "Joint Network Optimization and Downlink Beamforming for CoMP Transmissions using Mixed Integer Conic Programming," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 3972–3987, 2013.
- [10] Y. Teng and W. Zhao, "Robust Group Sparse Beamforming for Dense C-RANs with Probabilistic SINR Constraints," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017, pp. 1–6.
- [11] K. Zhang, W. Tan, G. Xu, C. Yin, W. Liu, and C. Li, "Joint RRH Activation and Robust Coordinated Beamforming for Massive MIMO Heterogeneous Cloud Radio Access Networks," *IEEE Access*, vol. 6, pp. 40 506–40 518, 2018.
- [12] Y. Shi, J. Zhang, and K. B. Letaief, "Robust group sparse beamforming for multicast green cloud-ran with imperfect csi," *IEEE Transactions on Signal Processing*, vol. 63, no. 17, pp. 4647–4659, 2015.
- [13] A. Zamani, O. Taghizadeh, G. Dartmann, R. Mathar, and A. Schmeink, "Joint User Association and Robust Beamforming Optimization for C-RANs with Wireless Fronthauls," in *IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2019.
- [14] A. J. Fehske, P. Marsch, and G. P. Fettweis, "Bit per joule efficiency of cooperating base stations in cellular networks," in *2010 IEEE Globecom Workshops*, 2010, pp. 1406–1411.
- [15] Q. Shi, M. Hong, X. Fu, and T. Chang, "Penalty Dual Decomposition Method for Nonsmooth Nonconvex Optimization," *CoRR*, vol. abs/1712.04767, 2017.
- [16] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*, 1st ed. Athena Scientific, 1996.
- [17] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [18] L. Gurobi Optimization, "Gurobi optimizer reference manual," 2020. [Online]. Available: <http://www.gurobi.com>
- [19] M. Razaviyayn, M. Hong, and Z. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, Aug. 2013.