Leveraging Muiti-cell NOMA for Cell Edge

Zhanwei Yu[1], Lei You[2] and Di Yuan[1]

[1]Department of Information Technology, Uppsala University, Sweden
[2]Independent Researcher, Sweden
Emails: zwyuapr1@gmail.com;lei.you@pm.me;di.yuan@it.uu.se

***Abstract*—Non-orthogonal multiple access (NOMA) is a promising technique for performance enhancement in cellular networks. This paper investigates how much NOMA can offer for cell edge in multi-cell scenarios. We consider the joint optimization problem of time-frequency resource allocation, user pairing, and power split, for scaling up the demand delivered for edge users. Our approach for problem solving consists in iteratively applying a fixed-point method to the cells, and, for each cell, deriving an algorithm guaranteeing optimum for single-cell optimization. By embedding the fixed-point method into bi-section search, we are able to show the overall approach guarantees global optimality. Numerical results demonstrate that multi-cell NOMA optimization has much more to offer over orthogonal multiple access for the experience of edge users, in particular for high-demand and resource-limited scenarios.**

***Index Terms*—NOMA, multi-cell, cell edge, resource allocation**

I. Introduction

Designing beyond-5G systems targets squeezing more bits out of the spectrum. Non-orthogonal multiple access (NOMA) is considered a promising radio access technique for performance enhancement. The basic idea of NOMA is using the same transmission channel to simultaneously serve multiple user equipments (UEs) at the cost of inter-user interference. By utilizing successive interference cancellation (SIC), NOMA can achieve better performance than orthogonal multiple access (OMA).

The experience of edge UEs is an important performance metric. There are a number of works studying improving this metric with NOMA. In [1], the performance of a single edge UE in a two-user multiple-input single-output (MISO) NOMA system is studied and three cooperative downlink transmission schemes with simultaneous wireless information and power transfer (SWIPT) are proposed. To improve the experience of the edge UEs in a two-user NOMA system, the authors of [2] propose two cooperative relaying schemes, called on/off full-duplex relaying and on/off half-duplex relaying. In [3], a two-cell NOMA system is considered, where two center UEs are served by their respective cells and one edge UE is served by both cells, the authors propose both centralized and distributed algorithms to minimize the total transmit power. In [4], the authors propose a downlink NOMA based coordinated direct and relay system with one center UE and multiple edge UEs, where a decode-and-forward and full-duplex relay is used to help the latter.

The works in [1]–[4] all focus on single-cell or two-cell scenarios. For more general multi-cell NOMA, resource optimization is more challenging, due to the interplay between SIC and inter-cell interference.

Recently, some studies have addressed general-scenario multi-cell NOMA. The authors of [5] consider the problem of dynamic power allocation with coordinated multi-point (CoMP) transmission in the downlink and propose a distributed power optimization approach. The authors of [6] study the cell load coupling system for multi-cell NOMA, and propose an algorithm for power allocation and UE pairing for optimal resource management. A complementary note to [6] is provided in [7]. The authors of [8] study the total transmit power minimization problem for multi-cell and multi-carrier NOMA (MCMC-NOMA) networks, and propose a resource allocation algorithm that can dramatically reduce the power consumption. The authors of [9] investigate energy optimization in MCMC-NOMA networks, and propose tailored algorithms to provide energy-efficient solutions for three NOMA grouping schemes.

Different from [1]–[4], which focus on single-cell or two-cell NOMA systems, we consider more general multi-cell scenarios. In addition, we focus on the experience of edge UEs. Indeed, the achievable rate of the edge UEs is generally much lower than that of the cell-center UEs [10]. We aim at investigating how much NOMA can offer to edge UEs. To this end, we study the joint optimization problem of time-frequency resource allocation, UE pairing, and power split, with the objective of maximizing the amount of data delivered to cell-edge UEs. We formulate this problem and derive its fundamental properties. We first propose a single-cell algorithm, aiming to minimize time-frequency resource consumption for any given amount of data to be delivered to the edge UEs within one cell, while the resource allocation of other cells is tentatively fixed. Next, we iteratively apply a fixed-point method with the single-cell algorithm to minimize multi-cell resource consumption. Then, we approach the global optimum of the overall problem by embedding the fixed-point method into bi-section search. Finally we evaluate the performance of the algorithm in multi-cell NOMA compared to OMA.

II. System Model and Problem Formulation

A. System Model

Denote by $\mathcal{I} = \{1, 2, ..., I\}$ and $\mathcal{J} = \{1, 2, ..., J\}$ the sets of cells and UEs, respectively. We consider downlink and denote by g_{ij} the channel power gain from cell i to UE j ($i \in \mathcal{I}, j \in \mathcal{J}$). Denote by $\mathcal{J}_i$ the set of UEs served by cell i.

The time-frequency resource is divided into resource units (RUs). In NOMA, multiple UEs can access the same RU by SIC. However, the complexity of decoding grows rapidly with the number of UEs involved in SIC. It has been demonstrated that two UEs in SIC can offer a good trade-off between complexity and performance [11]. In view of this, we consider two UEs in one cell sharing one RU as a *pair* (denoted by $u = \{j, h\}$), and call the process of selecting UEs to form pairs as *UE pairing* or simply *pairing*. To keep the generality, a pair may consist of a single UE j, i.e., UE uses the OMA mode, and we denote this special pair by $u = \{0, j\}$. For cell i, denote by $\mathcal{U}_i$ the set of all candidate pairs of UEs in $\mathcal{J}_i$. Similarly, use $\mathcal{U}_j$ to refer to the set of all pairs containing UE j. Furthermore, we use $\mathcal{U}$ to refer to the set of all candidate pairs, i.e., $\mathcal{U} = \bigcup_{i\in\mathcal{I}}\mathcal{U}_i$, and use N to denote the cardinality of U, i.e., $N = |\mathcal{U}|$. Also, we put indices on u to differentiate between the pairs, i.e., $\mathcal{U} = \{u_1, u_2, ..., u_N\}$. We use a binary y_u to indicate whether or not the pair u is selected. We optimize pairing by optimizing $\boldsymbol{y}$, where $\boldsymbol{y} = [y_{u_1}, y_{u_2}, ..., y_{u_N}]$.

We use p_i to denote the transmission power of cell i on an RU. For each pair u in cell i, *power splitting* is done on p_i, with q_{ju} and q_{hu} allocated to UE j and h, respectively. Namely,

$$p_i = q_{ju} + q_{hu}. \tag{1}$$

Besides, denote by x_u the proportion of RUs allocated to pair u. We have

$$\rho_i = \sum\nolimits_{u\in\mathcal{U}_i} x_u \leq \bar{\rho}, i \in \mathcal{I}, \tag{2}$$

where $\bar{\rho}$ is how many RUs at most can be allocated in one cell.

For any UE j ($j \in u$), the signal-to-interference and noise ratio (SINR) is computed by

$$\gamma_{ju} = \frac{q_{ju}g_{ij}}{B_{ju}q_{hu}g_{ij} + \sum_{k\in\mathcal{I}\setminus\{i\}} p_k g_{kj}\rho_k + \sigma^2}, j \in u, u \in \mathcal{U}. \tag{3}$$

We set $\gamma_{0u} = 0$ to for with the special case $u = \{0, j\}$.

In (3), $B_{ju}q_{hu}g_{ij}$ is the inter-user interference within a cell due to SIC, where B_{ju} is a binary indicator depending on j's decoding order. If UE j decodes h's signal first and hence h's signal does not affect j, $B_{ju} = 0$, otherwise, $B_{ju} = 1$. Besides, $\sum_{k\in\mathcal{I}\setminus\{i\}} p_k g_{kj}\rho_k$ is the inter-cell interference, where ρ_k is the amount of time-frequency resource consumption of cell k and reflects the likelihood that a UE outside cell k receives interference from k. The symbol σ^2 is the noise power.

In addition, in any pair u, the optimal decoding order (i.e., B_{ju} and B_{hu}) is determined by the channel condition, inter-cell interference, and noise [12]. We define w_j for the purpose of modeling the decoding order for any UE j as follows.

$$w_j = \left(\sum\nolimits_{k\in\mathcal{I}\setminus\{i\}} p_k g_{kj}\rho_k + \sigma^2\right) / g_{ij}. \tag{4}$$

In NOMA, the UE, which has better received signal in relation to inter-cell interference and noise, decodes the other UE in a pair first and there is no inter-user interference in its denominator of SINR. Thus, in any pair u, we have the following rule.

$$\begin{cases} B_{ju} = 0, B_{hu} = 1, \text{if } w_j < w_h, \\ B_{ju} = 1, B_{hu} = 0, \text{otherwise.} \end{cases} \tag{5}$$

For the special case of $u = \{0, j\}$, the value of w_0 is set to be a large value such that in effect UE j is in the OMA mode.

According to (3), we can obtain the achievable capacity for UE j in pair u on any RU as follows, where $\boldsymbol{q} = [q_{ju_1}, q_{hu_1}, q_{ju_2}, q_{hu_2}, ..., q_{ju_N}, q_{hu_N},]$, $\boldsymbol{\rho} = [\rho_1, \rho_2, ..., \rho_I]$, and $\boldsymbol{B} = [B_{ju_1}, B_{hu_1}, B_{ju_2}, B_{hu_2}, ..., B_{ju_N}, B_{hu_N},]$.

$$\begin{aligned} c_{ju}(\boldsymbol{q}, \boldsymbol{\rho}, \boldsymbol{B}) &= \log(1+\gamma_{ju}) = \log\left(1 + \frac{q_{ju}}{B_{ju}q_{hu} + w_j}\right) \\ &= \log\left(1 + \frac{q_{ju}g_{ij}}{B_{ju}q_{hu}g_{ij} + \sum_{k\in\mathcal{I}\setminus\{i\}} p_k g_{kj}\rho_k + \sigma^2}\right). \end{aligned} \tag{6}$$

We use d_j to denote the base demand of UE j ($j \in \mathcal{J}$) and denote by M and B the total number of RUs in one cell and the bandwidth of each RU, respectively. We have

$$\sum\nolimits_{u\in\mathcal{U}_j} MBc_{ju}(\boldsymbol{q}, \boldsymbol{\rho}, \boldsymbol{B})\, x_u \geq d_j, j \in \mathcal{J}. \tag{7}$$

In addition, we use $\mathcal{S}$ to denote the set of edge UEs. We would like to scale up the demand that can be delivered to the edge UEs by a *scaling factor* α ($\alpha \geq 1$) because considering the maximization of α can tell us how much NOMA can offer to edge UEs. Thus, we have

$$\sum\nolimits_{u\in\mathcal{U}_j} MBc_{ju}(\boldsymbol{q}, \boldsymbol{\rho}, \boldsymbol{B})\, x_u \geq \alpha d_j, j \in \mathcal{S}. \tag{8}$$

For convenience, in the following discussion, we use normalized d_j such that two notations M and B are not necessary.

B. Problem Formulation

Consider optimizing resource allocation in NOMA networks for demand scaling maximization of edge UEs. We optimize power split $\boldsymbol{q}$, pair-level resource allocation $\boldsymbol{x}$, pair selection $\boldsymbol{y}$, decoding order indicator $\boldsymbol{B}$, and cell-level resource allocation $\boldsymbol{\rho}$. The objective function α is the scaling factor for the given set of edge UEs $\mathcal{S}$ ($\mathcal{S} \subseteq \mathcal{J}$), with respect to the base UE demand $\boldsymbol{d} = [d_1, d_2, ..., d_m]$. The formulation is in (9).

$$[\textit{MaxD}] \quad \max_{\substack{\alpha\geq 1, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\rho}\geq \boldsymbol{0} \\ \boldsymbol{y}, \boldsymbol{B}\in\{0,1\}}} \alpha \tag{9a}$$

$$\text{s.t.} \sum\nolimits_{u\in\mathcal{U}_j} c_{ju}(\boldsymbol{q}, \boldsymbol{\rho}, \boldsymbol{B})\, x_u \geq \alpha d_j, j \in \mathcal{S} \tag{9b}$$

$$\sum\nolimits_{u\in\mathcal{U}_j} c_{ju}(\boldsymbol{q}, \boldsymbol{\rho}, \boldsymbol{B})\, x_u \geq d_j, j \in \mathcal{J}\setminus\mathcal{S} \tag{9c}$$

$$\sum\nolimits_{j\in u} q_{ju} \leq p_i, u \in \mathcal{U}_i, i \in \mathcal{I} \tag{9d}$$

$$\rho_i = \sum\nolimits_{u\in\mathcal{U}_i} x_u \leq \bar{\rho}, i \in \mathcal{I} \tag{9e}$$

$$x_u \leq y_u\bar{\rho}, u \in \mathcal{U} \tag{9f}$$

$$\sum\nolimits_{u\in\mathcal{U}_j} y_u \leq 1, j \in \mathcal{J} \tag{9g}$$

$$w_j = \frac{\left(\sum_{k\in\mathcal{I}\setminus\{i\}} p_k g_{kj}\rho_k + \sigma^2\right)}{g_{ij}}, j \in \mathcal{J}_i, i \in \mathcal{I} \tag{9h}$$

$$w_j > w_h \vee (w_j = w_h \wedge h < j) \Rightarrow B_{ju} = 1, \quad j \neq h, j, h \in u, u \in \mathcal{U}_i, i \in \mathcal{I} \tag{9i}$$

$$B_{ju}+B_{hu}=1, j\neq h, j,h\in u, u\in\mathcal{U}_i, i\in\mathcal{I} \quad (9j)$$

The UEs in $\mathcal{S}$ can be regarded as being throughput-oriented such that delivering more bits leads to higher satisfaction, as imposed by constraints (9b). The other UE d_j $(j\in\mathcal{J}\setminus\mathcal{S})$ constraints are (9c). Constraints (9d) and (9e) impose the cell power limit and RU limit of any cell, respectively. Constraints (9f) guarantee that RU allocation occurs only for selected pairs. By constraints (9g), each UE belongs up to one pair such that the selected pairs are not overlapping. Constraints (9h) - (9j) are for the decoding order in a pair, and they make decoding orders follow the rule (5) in Section II-A.

III. Fundamental Properties of MaxD

Before presenting the solution method for *MaxD*, we outline two lemmas for optimality characterization. The lemmas provide some fundamental properties for solving *MaxD*.

We denote by $\mathcal{H}$ a function that gives the normalized maximum time-frequency resource consumption of one cell, i.e.,

$$\mathcal{H}(\boldsymbol{\rho})=\frac{1}{\bar{\rho}}\max_{i\in\mathcal{I}}\rho_i. \quad (10)$$

Lemma 1. *If* $[\alpha^*,\boldsymbol{q}^*,\boldsymbol{x}^*,\boldsymbol{\rho}^*,\boldsymbol{y}^*,\boldsymbol{B}^*]$ *is optimal to MaxD, then* $\sum_{u\in\mathcal{U}_j}c_{ju}\left(\boldsymbol{q}^*,\boldsymbol{\rho}^*,\boldsymbol{B}^*\right)x_u^*=\alpha^*d_j$ *for some* $j\in\mathcal{S}$ *and* $\mathcal{H}(\boldsymbol{\rho}^*)=1$ *hold.*

Proof: For the first part of the lemma, consider an optimal solution $[\alpha^*,\boldsymbol{q}^*,\boldsymbol{x}^*,\boldsymbol{\rho}^*,\boldsymbol{y}^*,\boldsymbol{B}^*]$ to *MaxD*. Suppose this solution satisfies that all inequalities strictly hold in (9b). Then one can increase α^* to α' such that at least one of (9b) holds as equality. Thus, $[\alpha',\boldsymbol{q}^*,\boldsymbol{x}^*,\boldsymbol{\rho}^*,\boldsymbol{y}^*,\boldsymbol{B}^*]$ is a feasible solution and $\alpha'>\alpha^*$, which contradicts the assumption that $[\alpha^*,\boldsymbol{q}^*,\boldsymbol{x}^*,\boldsymbol{\rho}^*,\boldsymbol{y}^*,\boldsymbol{B}^*]$ is optimal. Therefore, there exists some j $(j\in\mathcal{S})$ such that $\sum_{u\in\mathcal{U}_j}c_{ju}\left(\boldsymbol{q}^*,\boldsymbol{\rho}^*,\boldsymbol{B}^*\right)x_u^*=\alpha^*d_j$.

For the second part of the lemma, obviously any solution with $\mathcal{H}(\boldsymbol{\rho})>1$ is infeasible because at least one of (9e) is violated. Now consider an optimal solution and suppose $\mathcal{H}^*(\boldsymbol{\rho})<1$. Suppose we increase $\boldsymbol{x}^*$ to $\boldsymbol{x}'=\beta\boldsymbol{x}^*$ $(\beta>1)$ such that $\boldsymbol{\rho}'=\beta\boldsymbol{\rho}^*$ satisfies $\mathcal{H}(\boldsymbol{\rho}')=1$. To show this is the case, we first define an auxiliary function $f(z)=(1+\frac{z}{\beta})^{\beta}$. According to Taylor's formula, the auxiliary function can be written as

$$f(z)=(1+\frac{z}{\beta})^{\beta}=1+z+R_1(z)>1+z, \quad (11)$$

where $R_1(z)=\frac{(\beta-1)(1+\frac{\theta z}{\beta})^{\beta-2}(\theta z)^2}{2\beta}>0$ $(\theta\in(0,1))$, which is the Lagrange remainder. For each UE j $(j\in\mathcal{S})$, with (12) we thus have

$$\sum_{u\in\mathcal{U}_j}c_{ju}\left(\boldsymbol{q}^*,\boldsymbol{\rho}',\boldsymbol{B}^*\right)x_u'=\sum_{u\in\mathcal{U}_j}c_{ju}\left(\boldsymbol{q}^*,\beta\boldsymbol{\rho}^*,\boldsymbol{B}^*\right)\beta x_u^*$$
$$=\sum_{u\in\mathcal{U}_j}\beta\log\left(1+\frac{q_{ju}^*g_{ij}}{B_{ju}^*q_{hu}^*g_{ij}+\beta\sum_{k\in\mathcal{I}\setminus\{i\}}p_kg_{kj}\rho_k^*+\sigma^2}\right)x_u^*$$
$$>\sum_{u\in\mathcal{U}_j}\log\left[1+\frac{q_{ju}^*g_{ij}}{\beta\left(B_{ju}^*q_{hu}^*g_{ij}+\sum_{k\in\mathcal{I}\setminus\{i\}}p_kg_{kj}\rho_k^*+\sigma^2\right)}\right]^{\beta}x_u^*$$
$$>\sum_{u\in\mathcal{U}_j}\log\left(1+\frac{q_{ju}^*g_{ij}}{B_{ju}^*q_{hu}^*g_{ij}+\sum_{k\in\mathcal{I}\setminus\{i\}}p_kg_{kj}\rho_k^*+\sigma^2}\right)x_u^*$$
$$=\sum_{u\in\mathcal{U}_j}c_{ju}\left(\boldsymbol{q}^*,\boldsymbol{\rho}^*,\boldsymbol{B}^*\right)x_u^*\geq\alpha d_j$$
$$\Rightarrow\sum_{u\in\mathcal{U}_j}c_{ju}\left(\boldsymbol{q}^*,\boldsymbol{\rho}',\boldsymbol{B}^*\right)x_u'>\alpha d_j \quad (12)$$

The same process applies for UE j $(j\in\mathcal{J}\setminus\mathcal{S})$. Therefore, $\boldsymbol{\rho}'$ is a feasible solution to *MaxD* such that all inequalities in (9b) and (9c) strictly hold. Under $\boldsymbol{\rho}'$, one can increase α^* to α' as earlier in the proof, to obtain a better objective value, which contradicts our assumption that $[\alpha^*,\boldsymbol{q}^*,\boldsymbol{x}^*,\boldsymbol{\rho}^*,\boldsymbol{y}^*,\boldsymbol{B}^*]$ is optimal. Thus, at the optimum of *MaxD* there is at least one UE j $(j\in\mathcal{S})$ such that $\sum_{u\in\mathcal{U}_j}c_{ju}\left(\boldsymbol{q}^*,\boldsymbol{\rho}^*,\boldsymbol{B}^*\right)x_u^*=\alpha^*d_j$ and $\mathcal{H}(\boldsymbol{\rho}^*)=1$. ■

By Lemma 1 we know that a solution is optimal to (9) only if there exists a cell that runs out time-frequency resource. Intuitively, if all cells have unused time-frequency resource, then one can improve the objective function such that more data would be delivered to UEs.

Lemma 2. $[\alpha^*,\boldsymbol{q}^*,\boldsymbol{x}^*,\boldsymbol{\rho}^*,\boldsymbol{y}^*,\boldsymbol{B}^*]$ *is optimal to MaxD if (9b) and (9c) all hold as equality and* $\mathcal{H}(\boldsymbol{\rho}^*)=1$.

Proof: Consider a solution satisfying the condition in the lemma, and $\mathcal{H}(\rho^*)=1$. Suppose the solution is non-optimal, i.e., there exists a better solution with $\alpha'>\alpha^*$. Replacing α^* by α' in (9b) causes (9b) to become violated. Thus x_u^* $(u\in\mathcal{U}_j\cup\mathcal{U}_i, j\in\mathcal{S})$ must increase to have (9b) remain satisfied. For any UE h $(h\in\mathcal{J}_k\setminus\mathcal{S}, k\neq i)$ in cell k $(k\neq i)$, the received inter-cell interference would grow, resulting in the violations of (9c). To satisfy (9c), x_u^* $(u\in\mathcal{U}_h\cap\mathcal{U}_k, h\in\mathcal{J}_k\setminus\mathcal{S}, k\neq i)$ must increase. In fact, all the x_u $(u\in\mathcal{U})$ must increase such that $\boldsymbol{\rho}^*$ must increase to $\boldsymbol{\rho}'$ for satisfying (9b) and (9c). Since $\mathcal{H}(\boldsymbol{\rho}*)=1$, we have $\mathcal{H}(\boldsymbol{\rho}')>1$ which contradicts the constraint (9e), i.e., infeasible. Hence he conclusion. ■

IV. Single-cell Optimization for Given Scaling Factor

Due to the interference among cells in multi-cell NOMA, one cell's pairing may affect the other cells' power splits, and vice versa. Enlightened by Lemma 1 and 2, we first study a subproblem of *MaxD* in this section: single-cell time-frequency resource consumption minimization problem with a given α and fixed $\boldsymbol{\rho}_{-i}$ (vector $\boldsymbol{\rho}_{-i}$ is composed of all elements but ρ_i of $\boldsymbol{\rho}$). In this special case, the variable $\boldsymbol{B}$ can be dropped, since $\boldsymbol{B}$ can be pre-determined according to (9h) - (9j). Then, we formulate this minimization problem as follows, where $\boldsymbol{q}_i,\boldsymbol{x}_i,\boldsymbol{y}_i$ is the corresponding variable elements for power split, pair-level resource allocation, and pairs selection.

$$\min_{\rho_i,\boldsymbol{q}_i,\boldsymbol{x}_i,\boldsymbol{y}_i,}\ \rho_i \text{ s.t. (9b)–(9g) of cell } i. \quad (13)$$

This minimization problem (14) is not straightforward even under fixed inter-cell interference, because this problem includes power split, pair-level resource allocation, and pairs selection. However, according to the following Theorem 3, we can draw a conclusion that the optimal power split is independent of pairs selection.

Theorem 3. *Consider* u $(u\in\mathcal{U}_i)$*. The optimal power split for given* $\hat{\boldsymbol{y}}_i$ *whose* $y_u=1$*, which is obtained by optimally* *solving (14), is also optimal for any other $\hat{\boldsymbol{y}}_i'$ whose $y_u = 1$.*

Proof: Denote by $\hat{\boldsymbol{q}}_u = [\hat{q}_{ju}, \hat{q}_{hu}]$ and $\hat{\boldsymbol{q}}_u' = [\hat{q}_{ju}', \hat{q}_{hu}']$ the optimal power split of pair u for $\hat{\boldsymbol{y}}_i$ and $\hat{\boldsymbol{y}}_i'$, respectively. Suppose $\hat{\boldsymbol{q}}_u$ is non-optimal for $\hat{\boldsymbol{y}}_i'$. Then there are two cases: 1) $c_{ju}(\hat{\boldsymbol{q}}_u) = c_{ju}(\hat{\boldsymbol{q}}_u')$ for any UE j in pair u; 2) $c_{ju}(\hat{\boldsymbol{q}}_u) \neq c_{ju}(\hat{\boldsymbol{q}}_u')$ for at least one UE j in pair u.

For 1), $\hat{\boldsymbol{q}}_u$ and $\hat{\boldsymbol{q}}_u'$ result in the same x_u for satisfying (9b) and (9c), and are equally good for (14) which contradicts our assumption. Thus $\hat{\boldsymbol{q}}_u$ is optimal for $\hat{\boldsymbol{y}}_i'$. For 2), assume $c_{ju}(\hat{\boldsymbol{q}}_u) > c_{ju}(\hat{\boldsymbol{q}}_u')\,(j \in u)$. By Lemma 2, $\hat{\boldsymbol{q}}_u'$ makes (9b) and (9c) become equalities under $\hat{\boldsymbol{y}}_i'$. Replacing $\hat{\boldsymbol{q}}_u$ by $\hat{\boldsymbol{q}}_u'$ leads to some slack in (9b) and (9c), hence the objective α can be improved. This contradicts that $\hat{\boldsymbol{q}}_u'$ is optimal for $\hat{\boldsymbol{y}}_i'$. The same proof applies to $c_{ju}(\hat{\boldsymbol{q}}_u) < c_{ju}(\hat{\boldsymbol{q}}_u')\,(j \in u)$. Hence the conclusion. ■

A. Finding Optimal Power Split

By Theorem 3, the optimal power split is decoupled from pairing. Then, we study how to find the optimal power split under fixed pairing in cell i.

With (9h) - (9j), the decoding order is fixed in any UE pair u. Without loss of generality, we assume UE j decodes first. Thus we have the rate of UEs j, h in UEs pair u in cell i are as follows.

$$c_{ju} = \log(1 + \frac{q_{ju}}{w_j}), \tag{14}$$

$$c_{hu} = \log(1 + \frac{q_{hu}}{q_{ju} + w_h}). \tag{15}$$

According to (15) and (16), we have

$$q_{ju} + q_{hu} = w_j e^{c_{ju} + c_{hu}} + (w_h - w_j) e^{c_{hu}} - w_h. \tag{16}$$

In addition, according to Lemma 2, c_{ju} and c_{hu} satisfies the following equations at any optimum.

$$c_{ju} = \mu_j d_j / x_u, \ c_{hu} = \mu_h d_h / x_u, \tag{17}$$

where $\mu_j, \mu_h = \alpha$ if UE $j, h \in \mathcal{S}$, otherwise $\mu_j, \mu_h = 1$. Then we define a function of x_u according to (15) - (18) as below.

$$f(x_u) = q_{ju} + q_{hu} = w_j e^{\frac{\mu_j d_j + \mu_h d_h}{x_u}} + (w_h - w_j) e^{\frac{\mu_h d_h}{x_u}} - w_h. \tag{18}$$

Considering RU power limit, the power split constraint reads

$$q_{ju} + q_{hu} \leq p_i. \tag{19}$$

Noted that constraint (21) hold as equality when x_u is minimized, because $f(x_u)$ is a monotonically decreasing function intuitively. That is,

$$w_j e^{\frac{\mu_j d_j + \mu_h d_h}{x_u}} + (w_h - w_j) e^{\frac{\mu_h d_h}{x_u}} - w_h = p_i. \tag{20}$$

Although the closed form of the minimized x_u is difficult to obtain for any given UE pair u, we can search for the minimized x_u by bisection search method according to (22). Then the corresponding optimal power split q_{ju}, q_{hu} can be obtained from (15) - (18). We remark that this processing of finding the optimal power split is also suitable for the case $u = \{0, j\}$, since we specially set the values of γ_{0u} and w_0.

Algorithm 1: Single-cell Optimization

Input: $\mathcal{S}$, $\boldsymbol{d}$, $\bar{\rho}$, $\boldsymbol{\rho}_{-i}$, α, $\mathcal{U}_i$
Output: $\rho_i^*, \boldsymbol{q}_i^*, \boldsymbol{x}_i^*, \boldsymbol{y}_i^*$;
1 **for** $u \in \mathcal{U}_i$ **do**
2   obtain $\min x_u$ of (22) by bisection search;
3     $W_u \leftarrow (T - \min\ x_u)$;
4 Construct $\mathcal{G}_i$ by (23);
5 $\mathcal{U}_i^* \leftarrow$ Maximum-Weighted-Matching($\mathcal{G}_i$);
6 $\boldsymbol{q}_i^* \leftarrow \boldsymbol{0}, \boldsymbol{x}_i^* \leftarrow \boldsymbol{0}, \boldsymbol{y}_i^* \leftarrow \boldsymbol{0}$,;
7 **for** $u \in \mathcal{U}_i^* \cap \mathcal{U}_i$ **do**
8     $y_u^* \leftarrow 1$, $x_u^* \leftarrow x_u$ from Step 2 ;
9     $q_{ju}^* \leftarrow q_{ju}, q_{hu}^* \leftarrow q_{hu}$ from (15) - (18);
10 $\rho_i^* \leftarrow \sum_{u \in \mathcal{U}_i} x_u^*$;
11 **return** $[\rho_i^*, \boldsymbol{q}_i^*, \boldsymbol{x}_i^*, \boldsymbol{y}_i^*]$;

B. Finding Optimal Pairing

By obtaining the minimized x_u for all u in cell i as shown earlier in Section IV-A, enumerating all pairings gives the optimal solution to (14). However, this exhaustive search does not scale, as the number of all candidate pairs $|\mathcal{U}_i|$ is exponential in the number of UEs. By the following derivation, we are able to obtain the optimum of (14) in polynomial time.

Theorem 4. *The optimal pairing of (14) is equivalent to the maximum weighted matching in the following undirected graph.*

$$\mathcal{G}_i = \begin{cases} \langle \mathcal{J}_i, \mathcal{U}_i \setminus \{\{0, j\} | j \in \mathcal{J}_i\}, \boldsymbol{W} \rangle & |\mathcal{J}_i| \text{ is even} \\ \langle \mathcal{J}_i \cup \{0\}, \mathcal{U}_i, \boldsymbol{W} \rangle & |\mathcal{J}_i| \text{ is odd.} \end{cases} \tag{21}$$

The graph is constructed by a 3-tuple, with the first element being the vertex set, the second element being the edge set, and the third element being the weight vector. The weight $\boldsymbol{W} = [W_{u_1}, W_{u_2}, ..., W_{u_N}]$ is calculated by

$$W_u = T - \min\ x_u\ (u \in \mathcal{U}_i), \tag{22}$$

where T is a positive value guarantees all weights being positive and min x_u is obtained from Section IV-A.

Proof: Without loss of generality, we first consider odd $|\mathcal{J}_i|$ case. By the definition in (23) each UE is corresponding to a vertex. For any pair $u = \{j, h\}$, there is one edge connecting the UE j and h, which represents j and h work in the NOMA mode. For any pair $u = \{0, j\}$, there is one edge connecting j and the 0, which represents UE j works in the OMA mode.

We remark that any $\boldsymbol{y}_i$ is feasible to (14) if and only if all the pair u with $y_u = 1$ $(u \in \mathcal{U}_i)$ form a matching in $\mathcal{G}_i$. Otherwise, there exist j such that $\sum_{u \in \mathcal{U}_j} y_u > 1$, and (9g) would be violated. Then by the definition of weights, minimizing ρ_i becomes finding a maximum weighted matching. All conclusions also hold for even $|\mathcal{J}_i|$ case. ■

The algorithm solving (14) is shown in Algorithm 1.

V. An Algorithmic Framework for MaxD

In this section, we present the algorithmic framework for deriving the optimum of *MaxD*. The algorithmic framework includes two parts: 1) a fixed-point method to solve multi-cell time-frequency resource consumption minimization with a given α; 2) by embedment of the fixed-point method in bisection search to find the optimal α^*.

A. Multi-cell Optimization for Given Scaling Factor

Recall that for the single-cell time-frequency resource consumption minimization problem with a given α, the optimum of (14) of cell i is a function of the time-frequency resource consumption of other cells $\boldsymbol{\rho}_{-i}$. We define it formally as follows

$$f_i\left(\boldsymbol{\rho}_{-i}\right) = \min_{\rho_i, \boldsymbol{q}_i, \boldsymbol{x}_i, \boldsymbol{y}_i,} \rho_i \text{ s.t. (9b)–(9g) of cell } i. \qquad (23)$$

It is proved that $f_i\left(\boldsymbol{\rho}_{-i}\right)$ is standard interference function (SIF) in [14]. Any SIF $f\left(\boldsymbol{\rho}\right)$ has the following two properties, where $\boldsymbol{\rho}, \boldsymbol{\rho}' \geq \mathbf{0}$.

- (Scalability) $\beta f\left(\boldsymbol{\rho}\right) > f\left(\beta \boldsymbol{\rho}\right)$ for any $\beta > 1$.
- (Monotonicity) $f\left(\boldsymbol{\rho}\right) \geq f\left(\boldsymbol{\rho}'\right)$ if $\boldsymbol{\rho} \geq \boldsymbol{\rho}'$.

We denote the optimal solution of (14) for the cells in $\mathcal{I}$ by $\boldsymbol{\rho}^*$. Based on the proportions of SIF, we can obtain the unique fixed point $\boldsymbol{\rho}^*$ with $\boldsymbol{\rho}^* = \boldsymbol{f}\left(\boldsymbol{\rho}^*\right)$ by fixed-point iterations on $\boldsymbol{f}$, where $\boldsymbol{f}\left(\boldsymbol{\rho}\right) = \left[f_1\left(\boldsymbol{\rho}_{-1}\right), f_n\left(\boldsymbol{\rho}_{-2}\right), ..., f_n\left(\boldsymbol{\rho}_{-n}\right)\right]$. Namely, for the iterative process $\boldsymbol{\rho}^{(k+1)} = \boldsymbol{f}\left(\boldsymbol{\rho}^{(k)}\right) (k \geq 0)$, we have $\lim_{k\to\infty} \boldsymbol{\rho}^{(k)} = \boldsymbol{\rho}^*$, for arbitrary non-negative staring point $\boldsymbol{\rho}^{(0)}$.

B. Finding Optimal Scaling Factor

By obtaining every cell's time-frequency resource consumption with a given α, the value of $\mathcal{H}(\boldsymbol{\rho})$ can be obtained also. With Lemma 1, we can drive the following corollary to find the optimal α^*.

Corollary 5. *For any given α and its corresponding minimized multi-cell time-frequency resource consumption $\boldsymbol{\rho}^*_\alpha$, $\alpha > \alpha^*$ if $\mathcal{H}(\boldsymbol{\rho}^*_\alpha) > 1$ and $\alpha < \alpha^*$ if $\mathcal{H}(\boldsymbol{\rho}^*_\alpha) < 1$.*

Thus, we update the value of α to approach the optimal α^* by bisection search according to Corollary 5. To be more specific, for the incumbent α the corresponding $\boldsymbol{\rho}^*_\alpha$ is obtained by Algorithm 1 and fixed-point iterations on $\boldsymbol{f}$. Then we calculate the corresponding $\mathcal{H}(\boldsymbol{\rho}^*_\alpha)$. According to Corollary 5, we know the optimal alpha is greater than α if $\mathcal{H}(\boldsymbol{\rho}^*_\alpha) < 1$, otherwise, the optimal alpha is smaller than α. That is, we search the optimal α^* in a given range $[\alpha_{\text{LB}}, \alpha_{\text{UB}}]$ by bisection search, leading to the maximum scaling factor α^* for $\mathcal{S}$. This algorithmic framework is detailed in Algorithm 2.

Algorithm 2: Scaling Factor Maximization

Input: $\mathcal{S}$, $\boldsymbol{d}$, $\bar{\rho}$, $\mathcal{U}$, $\epsilon > 0$
Output: $\alpha^*, \boldsymbol{q}^*, \boldsymbol{x}^*, \boldsymbol{\rho}^*, \boldsymbol{B}^*, \boldsymbol{y}^*$;

```
α_LB ← 1, α_UB ← 20;
repeat
    α ← (α_LB + α_UB)/2, k ← 0, ρ^(0) ← 0;
    repeat
        k ← k + 1, obtain ρ^(k) from Algorithm 1;
    until ‖ρ^(k) − ρ^(k−1)‖ < ε;
    if H(ρ^(k)) > 1 then
        α_UB ← α;
    else
        α_LB ← α;
until (α_UB − α_LB) < ε;
α* ← α, retrieve [q*, x*, ρ*, B*, y*] in Step 5 for α*;
return [α*, q*, x*, ρ*, B*, y*];
```

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Cell resource consumption limit $\bar{\rho}$	1.0
Total bandwidth	20 MHz
Path loss model	COST-231-HATA
Fading	Rayleigh flat fading
RB power in small cell	50 mW
RB power in macro cell	200 mW
Noise power spectral density σ^2	-174 dBm/Hz
Convergence tolerance ϵ	10^{-6}

VI. Performance Evaluation and Discussions

This section presents our results of performance evaluation. We use a specific topology to evaluate performance. In this topology, there are 21 cells including 7 macro cells and 14 small cells. Also, their gain and association matrices have been fixed. Reference demands are set to $\{1.0, 1.2, 1.4, 1.6, 1.8\}$ Mbps. Besides, we select the $P\%$ ($P = 10, 15, 25, 50, 75, 100$) UEs with worst gain in every cell into the set $\mathcal{S}$. Other parameters are given in Table I. The numerical results are illustrated in Figure 1, Figure 2 and Figure 3. We remark that both OMA and NOMA are solved to optimality.

Figure 1 illustrates the improvement of α (compared to $\alpha = 1.0$) with the increasing users uniform demand. Overall, the figure shows how scaling factor α decreases while the UE demand increases for four scenarios. It is can be easily explained that there is less time-frequency resource used for scaled UEs with increasing demand. In addition, NOMA can offer more to edge UEs compared to OMA. For example, the NOMA's scaling factor α are $\{2.2723, 1.75529, 1.35748\}$ under $d = \{0.07, 0.08, 0.09\}$, respectively. Compared to OMA (α are $\{1.9609, 1.46956, 1.09112\}$), NOMA increases 15.88%, 19.44% and 24.41% performance. It shows that NOMA performs better in the high-demand and resource-limited scenarios.

Figure 2 shows the scaling factor α with the increasing percentage of scaled UEs with $d = 0.07$. As we can see, the scaling factor α increases dramatically while the percentage of scaled UEs decreases. For example, when the percentage is equal to $\{10, 25, 50, 100\}\%$, the NOMA's and OMA's improvements of α are $\{4.37, 1.96, 1.56, 1.33\}$ and $\{5.47, 2.27, 1.75, 1.45\}$, respectively. Because as the percentage of scaled UEs decreases, time-frequency resource per scaled UEs increases. Each scaled UE can be allocated more resource. In addition, compared to OMA, NOMA increases

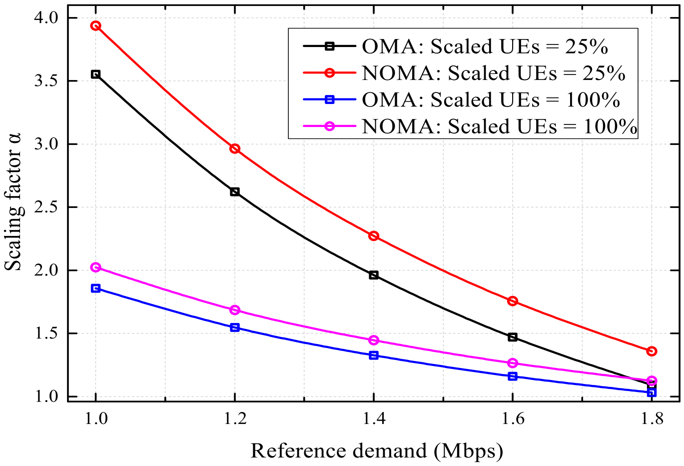

Fig. 1. This figure shows the scaling factor α in the function of UE demand at NOMA and OMA with different percentage of scaled UEs

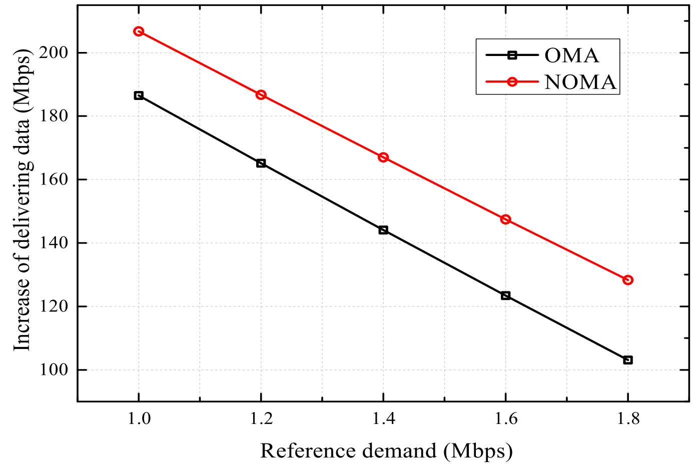

Fig. 3. This figure shows the increase of delivering data in the function of UE demand at NOMA and OMA with $P\% = 25\%$.

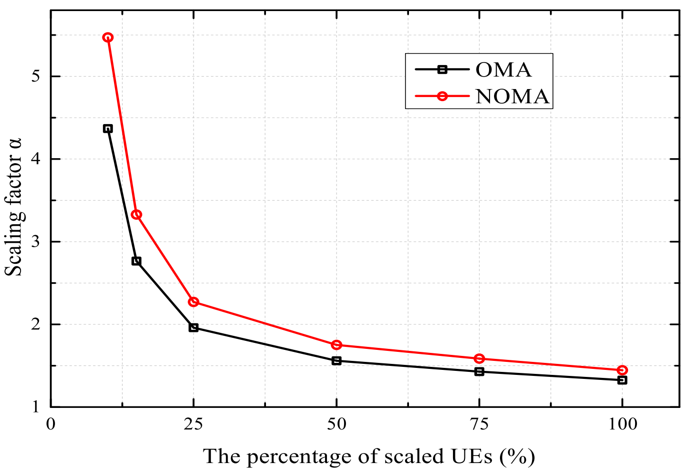

Fig. 2. This figure shows the scaling factor α in the function of the percentage of scaled UEs at NOMA and OMA with $d = 0.07$.

approximately 25% performance with $d = 0.07$. It shows that NOMA outperforms OMA on edge UEs performance.

As the scaling factor α is not intuitive for the network-wise performance, figure 3 shows the increase of delivering data in function of UE demand with $P\% = 25\%$. We can see that NOMA delivers more data ($\{10, 13, 15, 19, 24\}\%$, respectively) compared to OMA. That is, as resource more limited, NOMA can help edge UEs to deliver more data.

Besides, we find an interesting finding in our simulations, where the optimal pairing is always invariable in the process of searching α, regardless of how many α is. That is, the step 5 in Algorithm 1 is probably needed to implement only once during the whole process. We tend to believe that it is can be proved. If proved, the complexity of the algorithmic framework can be reduced greatly.

VII. Conclusions

In this paper, we have studied a joint optimization problem of time-frequency resource allocation, UEs pairing, and power split, aiming to improve the experience of edge UEs in multi-cell NOMA system. We have proposed an algorithmic framework for this problem, which can guarantee global optimality. Our numerical results have shown NOMA outperforms OMA in the field of improving the experience of edge UEs.

References

[1] T. N. Do, D. B. da Costa, T. Q. Duong and B. An, "Improving the performance of cell-edge users in MISO-NOMA systems using TAS and SWIPT-based cooperative transmissions," *IEEE Transactions on Green Communications and Networking*, vol 2, no. 1, pp. 49-62, 2017.

[2] T. N. Do, D. B. da Costa, T. Q. Duong and B. An, "Improving the performance of cell-edge users in NOMA systems using cooperative relaying," *IEEE Transactions on Communications*, vol. 66, no. 5, pp. 1883-1901, 2018.

[3] Q. Guo, C. W. Sung, Y. Chen, and C. S. Chen, "Power control for coordinated NOMA downlink with cell-edge users", *IEEE Wireless Communications and Networking Conference (WCNC)*, 2018, pp. 1-6.

[4] X. Pei, M. Wen and H. Yu, "NOMA-based coordinated direct and relay system with multiple cell-edge users", *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2019, pp. 1-5.

[5] M. S. Ali, E. Hossain, A. Al-Dweik, and D. I. Kim, "Downlink power allocation for CoMP-NOMA in multi-cell networks," *IEEE Transactions on Communications*, vol. 66, no. 9, pp. 3982-3998, 2018.

[6] L. You, D. Yuan, L. Lei, S. Sun, S. Chatzinotas, and B. Ottersten, "Resource optimization with load coupling in multi-cell NOMA," *IEEE Transactions on Wireless Communications*, vol. 17, no.7, pp. 4735-4749, 2018.

[7] L. You and D. Yuan, "A note on decoding order in optimizing multi-cell NOMA," *arXiv.org*, 2020. [Online]. Available: https://arxiv.org/abs/1909.08651.pdf

[8] D. Ni, L. Hao, Q. T. Tran and X. Qian, "Transmit power minimization for downlink multi-cell multi-carrier NOMA networks," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2459-2462, 2018.

[9] L. Lei, L. You, Y. Yang, D. Yuan, S. Chatzinotas and B. Ottersten, "Load coupling and energy optimization in multi-cell and multi-carrier NOMA networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp.11323-11337, 2019.

[10] L. Dai, B. Wang, Y. Yuan, S. Han, C. l. I, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74-81, 2015.

[11] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 721-742, 2017.

[12] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge university press, 2005.

[13] "IEC 80000-13:2008, quantities and units part 13: Information science and technology," International Electrotechnical Commission, 2008.

[14] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341-1347, 1995.