Computer Communications xxx (xxxx) xxx

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

# Resource allocation for downlink non-orthogonal multiple access in joint transmission coordinated multi-point networks

Mohamad Khattar Awad [a,*], Mohammed W. Baidas [b], Ahmad A. El-Amine [a]

[a] Department of Computer Engineering, College of Engineering and Petroleum, Kuwait University, Kuwait
[b] Department of Electrical Engineering, College of Engineering and Petroleum, Kuwait University, Kuwait

## ABSTRACT

Joint transmission coordinated multi-point (JT-CoMP) and non-orthogonal multiple access (NOMA) are key enabling technologies of 5G ubiquitous broadband infrastructures. These technologies are jointly expected to exploit multi-cell and non-orthogonal resource transmissions; thus, conventional resource allocation schemes that only consider either one of them fail to efficiently exploit resources of 5G networks. In this paper, we bridge this gap by proposing a practical and comprehensive joint sub-carrier assignment and power allocation scheme for network sum-rate maximization in JT-CoMP-enabled NOMA networks. We formulate the problem as a mixed integer non-linear programming (MINLP) problem, which is NP-hard. The problem is decoupled into two sub-problems, where the sub-carrier assignment is modeled as a two-sided many-to-many matching game and the power allocation is formulated as a difference of convex (DC) programming problem. The matching algorithm is proved to converge to a two-sided exchange stable matching. Furthermore, the solution computed by the proposed scheme is verified against a baseline solution computed by a commercial optimization package, and has been shown to achieve 91.38% of the baseline solution for JT-CoMP-NOMA networks. Simulation results illustrate that the proposed scheme enhances cell-edge users' achievable rates by $0.1 - 27.7\%$ in JT-CoMP-NOMA over conventional NOMA.

## 1. Introduction

Non-orthogonal multiple-access (NOMA) has recently been proposed as a promising radio access technique for the upcoming fifth generation (5G) cellular networks, which is due to its ability to improve spectral efficiency, fairness and massive connectivity, while reducing latency and signaling costs [1–4]. In NOMA networks, multiple users' signals are simultaneously multiplexed in the power-domain, such that successive interference cancellation (SIC) can be effectively applied to exploit performance and capacity gains [5]. However, transmitting all users' signals over a single channel, as in single-carrier NOMA, introduces unnecessary SIC complexity and decoding delay because the user with the highest channel gain will have to decode all other users' signals before decoding its own signal [1,6]. Thus, unlike single-carrier NOMA, multi-carrier NOMA reduces the system complexity by dividing the users into groups – which are not necessarily mutually exclusive – and multiplexing each group on different orthogonal sub-carriers (SCs) [1].

Multi-cell NOMA networks pose more challenges than single-cell NOMA due to the interplay and mutual influence of intra- and inter-cell interferences, and the lack of optimal capacity-achieving transmit/receive strategies [7]. Several recent research works have focused on resource allocation in downlink multi-cell NOMA networks. For example, the authors in [8] studied the problems of sum-power minimization (SPM) and sum-rate maximization (SRM) for downlink multi-cell NOMA networks. Particularly, in the SPM problem, closed-form solutions that can be optimally solved via the standard interference function are obtained. It was demonstrated that the derived solutions are superior to those of orthogonal frequency division multiple-access (OFDMA); however, at the expense of increased control overhead and computational complexity. In [9], the total transmit power minimization of the network base-stations (BSs) subject to users' rate requirements in downlink multi-cell NOMA networks is considered. It was shown that the proposed distributed power control algorithm significantly outperforms orthogonal multiple access (OMA) in terms of power consumption and outage probability.

Beside the resource allocation works proposed above, several multi-cell NOMA solutions have been proposed to manage inter-cell interference (ICI) with coordinated scheduling or beamforming [7]. For instance, in NOMA with joint-transmission (JT-NOMA), multiple BSs simultaneously serve cell-edge users at the expense of slightly reduced rates for cell-center users. Another solution proposes the use of NOMA with dynamic cell selection (DCS-NOMA), where a user's data is shared among multiple BSs, but is served by a single BS. In NOMA with coordinated scheduling (CS-NOMA), BSs coordinate to serve users such that ICI is reduced and quality-of-service (QoS) is guaranteed for cell-edge users. Coordinated beamforming with NOMA (CB-NOMA) is proposed to allow BSs to optimize their beamforming vectors so as to eliminate ICI and improve data rates of cell-edge users.

Coordinated multi-point (CoMP) transmission has been proposed by the third generation partnership project (3GPP), which was aimed

* Corresponding author.
*E-mail addresses:* mohamad@ieee.org (M.K. Awad), baidas@ieee.org (M.W. Baidas), ahmad.elamine@ieee.org (A.A. El-Amine).

https://doi.org/10.1016/j.comcom.2021.03.025
Received 25 July 2020; Received in revised form 24 December 2020; Accepted 27 March 2021
Available online xxxx
0140-3664/© 2021 Elsevier B.V. All rights reserved.

Please cite this article as: M.K. Awad, M.W. Baidas and A.A. El-Amine, Resource allocation for downlink non-orthogonal multiple access in joint transmission coordinated multi-point networks, Computer Communications (2021), https://doi.org/10.1016/j.comcom.2021.03.025.

at coordinating joint transmissions of BSs to users susceptible to ICI, i.e., cell-edge users, in the fourth generation (4G) cellular networks and beyond [10]. This has triggered several studies investigating resource allocation for CoMP-enabled NOMA networks. For instance, a general framework with distributed power allocation for CoMP-NOMA in downlink multi-cell networks is outlined in [11]. Particularly, the authors investigate the necessary conditions for the applicability of different CoMP schemes along with different users' spatial distributions. More importantly, it has been demonstrated that CoMP-NOMA yields significant spectral efficiency gains over CoMP-enabled OMA networks (CoMP-OMA). In [12], the authors study the problem of total transmit power minimization subject to target rate constraints in a two-cell CoMP-NOMA network, where two users are located near their associated BSs and a single cell-edge user is served by both BSs. A centralized optimal and distributed sub-optimal solutions were proposed, which yielded superior improvement over OMA in terms of power consumption and outage probability. Dynamic power allocation for sum-rate maximization in downlink multi-cell joint-transmission (JT)-CoMP-NOMA networks is considered in [13]. For these networks, a distributed power allocation optimization approach is proposed, whereby each BS can determine its optimal power allocation solution independently of the other coordinating BSs. It was demonstrated that the developed power allocation solution for JT-CoMP-NOMA yields significant spectral and energy efficiency gains in comparison to JT-CoMP-OMA. In [14], the joint design of power allocation and beamforming is considered for multi-cell multiuser multiple-input multiple-output (MIMO) COMP-enabled NOMA networks. Results reveal that significant performance gains can be achieved when grouping users based on their QoS requirements rather than their channel gains. In [15], the authors generalize the JT-COMP NOMA transmission to all UEs in the network, i.e., both cell-edge and cell-center, and developed a power minimization scheme to assign power to cooperating basestations under minimum UEs rate constraints. It was demonstrated – based on closed-form approximations of the outage probability and outage capacity – that spectral efficiency for NOMA clusters with large number of UEs can be improved by increasing the number of BSs. A joint access points clustering and power minimization for COMP-enabled NOMA green Internet-of-Things (IoT) networks is developed in [16]. Results reported in [16] demonstrate that the JT-COMP mitigates interference effectively to save transmit power of IoT access points while maintaining minimum rate requirements of IoT devices.

### 1.1. Motivation and contributions

Matching theory has recently gained significant attention from researchers in wireless communications due to its ability to provide tractable and low-complexity mathematical solutions for combinatorial problems, where matching between players in two distinct sets is required [17]. For instance, in [18], the resource allocation problem in downlink heterogeneous networks (HetNets) is studied, where small-cell BSs (SBSs) reuse the resource blocks (RBs) of the macro-cell BS (MBS) in order to efficiently serve users and thus further improve the spectral efficiency and resource utilization. For assigning RBs to SBSs, the authors model the problem as a many-to-one matching game with peer effects between SBSs and RBs. To ensure network stability and further improve the network sum-rate, swap operations are employed. Power control for each SBS is then solved by adopting sequential convex programming (SCP). In [19], the problem of coordinated device-to-device (D2D) communications underlaying NOMA cellular networks is considered. A two-stage matching game is modeled to solve the sub-channel assignment problem. In the first stage, a many-to-one matching game with externalities is developed to assign sub-channels to cellular users (CUs). In the second stage, another matching game is developed in order to assign sub-channels to D2D pairs for spectrum reuse, where each D2D pair can occupy a single sub-channel, while each sub-channel can be reused by as many D2D pairs as long as interference protection can be guaranteed for the CUs in that sub-channel. In [20], the authors consider the problem of joint sub-carrier assignment and power allocation in single-cell NOMA networks, where a many-to-many matching game with externalities is utilized to solve the problem of sub-carrier assignment, while geometric programming (GP) is used to solve the power allocation problem. In [21], the problem of joint subcarrier assignment and global energy-efficient power allocation (J-SA-GEE-PA) for energy harvesting two-tier downlink NOMA HetNets is considered. The authors aim to solve the J-SA-GEE-PA while maximizing the global energy-efficiency (GEE). Due to the non-convexity and NP-hardness of the formulated problem, the authors decouple the problem into two sub-problems: (1) many-to-many matching game using Gale–Shapley algorithm for subcarrier assignment, and (2) low-complexity algorithm to optimally solve the GEE power allocation problem.

The problem of resource allocation in JT-CoMP-NOMA networks has received less attention from the research sector, where most of the research on NOMA has considered single-cell scenarios with a limited number of practical constraints taken into account. For instance, user pairing in NOMA networks is a very critical task since users sharing the same spectrum resources are multiplexed in the power-domain and inter-user interference is thus introduced [22–24]. Therefore, careful user pairing should be considered in order to ensure successful SIC decoding. Dynamic power allocation is also crucial to realize the performance gains of NOMA networks. However, most of the existing research on JT-CoMP-NOMA and conventional NOMA (C-NOMA) either consider user pairing with a fixed power allocation strategy or consider dynamic power allocation without taking into account the user pairing effects [8,9,11–16,25]. In addition, most of the existing research on JT-CoMP-NOMA assume SCs are assigned to users, and focus solely on the power allocation problem [26,27].

In this paper, we examine and formulate the network sum-rate maximization for joint sub-carrier and power allocation problem in JT-CoMP-NOMA networks. Furthermore, cell-edge users receive joint transmissions from coordinated neighboring BSs[1] in order to enhance their data rates by mitigating inter-cell interference. Practical constraints, such as maximum number of users multiplexed over each SC, SIC decoding order, user pairing, and minimum rate requirements, are considered. The formulated problem is NP-hard due to the existence of inter- and intra-cell interferences [28]. The problem is then decomposed into two sub-problems, namely a SC assignment problem and a power allocation problem. We pose the SC assignment problem as a two-sided many-to-many stable matching game with *externalities* to assign the users to the SCs in the multi-cell network, while accounting for intra- and inter-cell interferences. Specifically, due to the existence of externalities, we enable *swap operations* to ensure stability. In addition, to solve the user pairing problem, we model the relationship between users in the network using an *undirected graph*. Then, a *cliques* formation algorithm [29] is employed to select the group of users that satisfy the user pairing requirements and form a NOMA cluster. The non-convex power allocation problem is modeled as a difference of convex (DC) programming problem [30], and solved via the Frank–Wolfe algorithm [31]. We show that the proposed algorithm converges to a local optimal solution in a finite number of iterations. The contributions of this work can be summarized as follows:

(1) Modeled the network sum-rate maximization problem, subject to all network practical constraints, as a mixed-integer non-linear programming (MINLP) problem, which is non-convex and combinatorial in nature.
(2) Decoupled the problem into a SC assignment problem and a power allocation problem, and proposed a novel resource allocation scheme for JT-CoMP-NOMA networks.

[1] In JT-CoMP networks, the coordinated BSs simultaneously transmit the same signal to a cell-edge user receiving joint transmissions [11].

**Table 1**
Notation.

| Symbols | Definition |
|---|---|
| $\mathcal{B}, \mathcal{U}, \mathcal{K}$ | Set of BSs, UEs and SCs, respectively |
| $\mathcal{T}, \mathcal{E}$ | Set of cell-center and cell-edge UEs, respectively |
| $\mathcal{C}_k^b$ | Set of UEs forming a NOMA cluster |
| $\mathcal{S}_u$ | Set of serving BSs of UE $u$ |
| $B, U, K$ | Number of BSs, UEs and SCs, respectively |
| $T, E$ | Number of cell-center and cell-edge UEs, respectively |
| $\gamma_{u,k}^b$ | Normalized channel gain from BS $b$ to UE $u$ over SC $k$ |
| $\bar{P}^b$ | Total transmit power of BS $b$ |
| $p_{u,k}^b$ | Power allocated to UE $u$ in cell $b$ over SC $k$ |
| $I_{u,k}, \varphi_{u,k}$ | Intra- and inter-cell interferences of UE $u$ over SC $k$, respectively |
| $\pi_k^b(u)$ | Decoding order of UE $u$ in cluster $\mathcal{C}_k^b$ |
| $\kappa$ | Maximum number of SCs assigned to each UE |
| $\mu$ | Maximum number of UEs multiplexed over each SC |
| $\tau$ | Number of coordinated BSs |
| $\Psi$ | Many-to-many matching between UEs and SCs |
| $\Psi(u), \Psi(k)$ | Matchings of UE $u$ and SC $k$ in $\Psi$, respectively |
| $\Psi_{uk'}^{uk}$ | Swap matching |
| $U_k(\mathcal{N})$ | Utility of SC $k$ over a subset of UEs $\mathcal{N}$ |
| $U_u(k)$ | Utility of UE $u$ over SC $k$ |
| $U_i(\Psi)$ | Utility of player $i$ in matching $\Psi$ |
| $\mathcal{P}_u, \mathcal{P}_k$ | Preference lists of UEs and SCs |
| $\mathbb{G}_k^b$ | UE pairing graph |
| $\mathbb{V}_k^b, \mathbb{L}_k^b$ | Vertices and undirected edges or links of $\mathbb{G}_k^b$, respectively |
| $\mathbb{S}_k^b$ | Complete sub-graph of $\mathbb{G}_k^b$ |

(3) Demonstrated that the proposed resource allocation scheme is generic, and thus can be used for JT-CoMP-NOMA, C-NOMA and OFDMA networks.
(4) Demonstrated that JT-CoMP-enabled NOMA networks are able to enhance the data rates of cell-edge UEs in comparison to C-NOMA.
(5) Analyzed the properties of the proposed scheme, and illustrated that it can be executed efficiently and with low computational complexity.
(6) Compared the solutions of the proposed scheme to baseline solutions computed by a commercial package [32], and illustrated that the proposed scheme provides a comparable network sum-rate performance.

To the best of our knowledge, no prior work has considered the problem of joint sub-carrier assignment and power allocation for network sum-rate maximization in JT-CoMP-NOMA networks, under a variety of practical constraints.

### 1.2. Paper organization and notation

The remainder of this paper is organized as follows. Section 2 presents the network model. Section 3 outlines the joint sub-carrier assignment and power allocation problem formulation. Section 4 discusses the many-to-many matching game for sub-carrier assignment, and Section 5 presents the power allocation problem formulation along with the proposed Frank–Wolfe algorithm. The proposed joint sub-carrier assignment and power allocation scheme is presented in Section 6. The simulation results are given in Section 7, and conclusions are drawn in Section 8. The frequently used symbols in this paper are summarized in Table 1.

## 2. Network model

We consider a downlink multi-cell NOMA network that supports JT-CoMP communications. The network consists of $B$ BSs denoted by the set $\mathcal{B} = \{1, \ldots, B\}$ and $U$ user equipments (UEs) denoted by the set $\mathcal{U} = \{1, \ldots, U\}$. The frequency spectrum allocated to a BS $b$ is divided into $K^b$ SCs and denoted by the set $\mathcal{K}^b = \{1, \ldots, K^b\}$. Note that in JT-CoMP systems, all coordinated BSs transmit on the same SC, i.e., the frequency reuse factor is equal to 1, i.e., $\mathcal{K}^b = \mathcal{K}, \forall b \in \mathcal{B}$.[2] The set of UEs associated with BS $b$ is denoted by $\mathcal{U}^b$. The channel state information (CSI) is perfectly known to the BSs, and the channel coefficient from BS $b$ to UE $u$ on SC $k$ is denoted by $h_{u,k}^b = g_{u,k}^b \left(d_u^b\right)^{-\alpha/2}$, where $g_{u,k}^b$ is an i.i.d. circularly symmetric complex Gaussian random variable representing the Rayleigh fading, $d_u^b$ is the propagation distance between BS $b$ and UE $u$, and $\alpha$ is the path-loss exponent. The additive white Gaussian noise (AWGN) at UE $u$ is denoted by $n_u$, which is zero-mean with variance $\sigma_u^2 = \sigma^2, \forall u \in \mathcal{U}$. Therefore, the corresponding normalized channel gain of UE $u$ is defined as $\gamma_{u,k}^b = \frac{|h_{u,k}^b|^2}{\sigma^2}$.

In JT-CoMP systems, UEs are categorized either as cell-edge users or as cell-center users based on comparing their normalized average channel gains from the BSs in their range to a threshold $\gamma^{thr}$ [13,25]. A UE $u$ with average channel gain satisfying the condition,

$$\frac{1}{K^b} \sum_{k \in \mathcal{K}^b} \gamma_{u,k}^b \leq \gamma^{thr}, \quad \forall b \in \mathcal{B}, \forall u \in \mathcal{U}^b, \tag{1}$$

in a given cell $b$ is considered a cell-edge UE in that particular cell; otherwise, it is considered as a cell-center UE. The threshold can be computed based on a number of network parameters like cell-edge SNR, number of users, backhaul load, and inter-cell interference. The design and optimization of the threshold is beyond the scope of this work; however, several methods have been proposed in the literature to optimize the threshold as in [33–36]. The sets of cell-center and cell-edge users with respect to BS $b$ are denoted by $\mathcal{T}^b$ and $\mathcal{E}^b$, respectively. Note that, a UE can either be a cell-edge or cell-center in a given cell, i.e., $\mathcal{T}^b \cap \mathcal{E}^b = \emptyset$. Consequently, the sets of all cell-center UEs and cell-edge UEs in the network are given by $\mathcal{T} = \bigcup_{b \in \mathcal{B}} \mathcal{T}^b$ and $\mathcal{E} = \bigcup_{b \in \mathcal{B}} \mathcal{E}^b$, respectively. In addition, the set of users receiving JT-CoMP transmissions from both BSs $b$ and $b'$, where $b \neq b'$, on SC $k$ is denoted by $\mathcal{O}_k^{b,b'}$.

NOMA networks exploit users channel gain differences and multiplex users in the power domain over the same SC [13]. Therefore, unlike OMA systems, a SC $k$ in cell $b$ can be allocated to multiple UEs, forming a NOMA cluster denoted by $\mathcal{C}_k^b$. Moreover, each UE $u$ can receive signals from the BS over multiple SCs [5,20]. The power allocated to UE $u$ in cell $b$ over SC $k$ is denoted by $p_{u,k}^b$. The allocated power to all UEs over the available spectrum in a given cell must satisfy the BS power budget constraint as follows,

$$\sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} p_{u,k}^b \leq \bar{P}^b, \quad \forall b \in \mathcal{B}, \tag{2}$$

where $\bar{P}^b$ is the total available power budget of BS $b$.

A cell-center UE, $u \in \mathcal{T}^b$, receives a superposed signal of all UEs signals in the NOMA cluster, expressed as

$$y_{u,k}^b = h_{u,k}^b \sum_{u' \in \mathcal{C}_k^b} \sqrt{p_{u',k}^b}\, x_{u',k}^b + \chi_{u,k}^b + n_{u,k}^b, \tag{3}$$

where $x_{u',k}^b$ is the signal of UE $u'$ assigned to cluster $\mathcal{C}_k^b$. $\chi_{u,k}^b$ denotes the inter-cell interference received from neighboring BSs transmitting on SC $k$, is given by

$$\chi_{u,k} = \sum_{b' \in \{\mathcal{B} \setminus b\}} h_{u,k}^{b'} \sum_{j \in \mathcal{C}_k^{b'}} \sqrt{p_{j,k}^{b'}}\, x_{j,k}^{b'}. \tag{4}$$

Furthermore, $n_{u,k}^b$ is the received AWGN. Unlike cell-center UEs, the signal received by a cell-edge UE $u \in \mathcal{E}^b$ from its coordinated set of serving BSs[3], denoted by $\mathcal{S}_u$, is given by [26]

$$y_{u,k}^b = \sum_{b \in \mathcal{S}_u} h_{u,k}^b \sum_{u' \in \mathcal{C}_k^b} \sqrt{p_{u',k}^b}\, x_{u',k}^b + \phi_{u,k} + n_{u,k}^b, \tag{5}$$

[2] In this paper, the SC sets $\mathcal{K}$ and $\mathcal{K}^b$ are used interchangeably.
[3] Note that the set of serving BSs for either a cell-center UE $u \in \mathcal{T}^b$ or a cell-edge UE $u \in \mathcal{E}$ that is not receiving JT-CoMP transmissions consists of the sole serving BS, i.e., $|\mathcal{S}_u| = 1$.

where $\phi_{u,k}$ denotes the inter-cell interference received from BSs that are not in the coordinated set of $u$, and yet transmitting on SC $k$, is given by

$$\phi_{u,k} = \sum_{b' \in \{\mathcal{B} \setminus S_u\}} h^{b'}_{u,k} \sum_{j \in C^{b'}_k} \sqrt{p^{b'}_{j,k}}\, x^{b'}_{j,k}. \tag{6}$$

It is worthwhile to note that a cell-edge UE does not necessarily receive a JT-CoMP transmission. Consequently, the cardinality of the set of coordinated BSs for cell-edge UE $u$ in this case is 1, i.e., $|S_u| = 1$.

Since a SC $k \in \mathcal{K}$ is assigned to a given NOMA cluster $C^b_k$, the signal of any UE $u \in C^b_k$ causes interference to other UEs $u' \in C^b_k$. To decode the intended message, each UE $u$ performs successive interference cancellation (SIC) on the superposed signal. Let $\pi^b_k(u)$ be the order of UE $u$ in a NOMA cluster $C^b_k$ based on an increasing order of their channel gains as follows

$$\gamma^b_{u^1,k} < \gamma^b_{u^2,k} < \cdots < \gamma^b_{u^C,k}, \quad \forall u \in C^b_k \text{ and } C = |C^b_k|. \tag{7}$$

A UE $u \in C^b_k$ successively decodes the signals of other UEs $j \in \left\{C^b_k | \pi^b_k(j) < \pi^b_k(u)\right\}$ with lower channel gains and cancels them before it decodes its desired signal. However, signals of UEs $j' \in \left\{C^b_k | \pi^b_k(j') > \pi^b_k(u)\right\}$ with larger channel gains are treated as intra-cell interference [20]. Therefore, according to SIC principle [37,38], UEs with lower channel gains are allocated higher power levels, whereas UEs with higher channel gains are allocated lower power levels [22]. Therefore, the following condition must hold for all UEs in a cluster $C^b_k$

$$p^b_{u,k} \geq p^b_{j,k}, \quad \forall u, j \in \left\{C^b_k | \pi^b_k(u) < \pi^b_k(j)\right\}, u \neq j. \tag{8}$$

Let $\eta^b_{u,k}$ be a binary decision variable, such that

$$\eta^b_{u,k} = \begin{cases} 1, & \text{if UE } u \text{ is assigned to SC } k \text{ in cell } b, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

The achievable data rate, in bps/Hz, for any UE $u \in \mathcal{U}$ served by the set of BSs, $S_u$, over SC $k \in \mathcal{K}$ is given by [11,20,26],

$$R_{u,k} = \log_2\left(1 + \frac{\sum_{b \in S_u} \eta^b_{u,k} p^b_{u,k} \gamma^b_{u,k}}{1 + I_{u,k} + \varphi_{u,k}}\right), \tag{10}$$

where $I_{u,k}$ and $\varphi_{u,k}$ denote the intra- and inter-cell interference terms, respectively, and are defined as follows

$$I_{u,k} = \sum_{b \in S_u} \gamma^b_{u,k} \sum_{i \in \{C^b_k | \pi^b_k(i) > \pi^b_k(u)\}} p^b_{i,k}, \tag{11}$$

and

$$\varphi_{u,k} = \sum_{b \in \{\mathcal{B} \setminus S_u\}} \gamma^b_{u,k} \sum_{j \in C^b_k} p^b_{j,k}. \tag{12}$$

Note that for a cell-center UE $u \in \mathcal{T}^b$, the achievable data rate, intra-cell interference and inter-cell interference reduce to $R_{u,k} = \log_2\left(1 + \frac{\eta^b_{u,k} p^b_{u,k} \gamma^b_{u,k}}{1 + I_{u,k} + \chi^b_{u,k}}\right)$, $I_{u,k} = \gamma^b_{u,k} \sum_{i \in \{C^b_k | \pi^b_k(i) > \pi^b_k(u)\}} p^b_{i,k}$, and $\chi^b_{u,k}$, respectively, since $S_u = b$. To guarantee quality of service (QoS) for all UEs on their assigned SCs[4], $R_{u,k}$ must be at least $R_{min}$, i.e.,

$$R_{u,k} \geq R_{min}. \tag{13}$$

In order to successfully perform SIC,[5] UEs that are scheduled in a NOMA cluster must have distinct channel gains [22–24]. That is, the channel gain difference between any two UEs $u, u' \in C^b_k, u \neq u'$, must be greater than threshold $\nu^b_k$.

$$|\gamma^b_{u,k} - \gamma^b_{u',k}| > \nu^b_k, \quad \forall b \in \mathcal{B}, \forall k \in \mathcal{K}, \forall u, u' \in C^b_k. \tag{14}$$

[4] Note that the QoS is defined over each SC, which is commonly adopted in literature to simplify the resource allocation design [39–41], and to ensure successful SIC on each assigned SC. In addition, this ensures efficient utilization of the scarce channel resources.

[5] Perfect SIC is assumed in this work.

Considering the complexity of SIC, which is in the order of $O\left(|C^b_k|^3\right)$ [20,42], the number of UEs multiplexed over one SC is limited to $\kappa$. Furthermore, the number of SCs assigned to a given UE is limited to $\mu$. Therefore, the following constraints must be satisfied,

$$\sum_{u \in \mathcal{U}} \eta^b_{u,k} \leq \kappa \text{ and } \sum_{k \in \mathcal{K}} \eta^b_{u,k} \leq \mu, \ \forall b \in \mathcal{B}, \forall k \in \mathcal{K}, \forall u \in \mathcal{U}. \tag{15}$$

Furthermore, considering the JT latency and overhead between coordinated BSs, the number of serving BSs for a cell-edge UE $u \in \mathcal{E}$ is limited to $\tau$. On the other hand, the maximum number of serving BSs for a cell-center UE, $u \in \mathcal{T}$, is 1. Thus, the following conditions arise,

$$|S_u| \leq \tau, \ \forall u \in \mathcal{E} \quad \text{and} \quad |S_u| \leq 1, \ \forall u \in \mathcal{T}. \tag{16}$$

Because UEs are scheduled on non-orthogonal spectrum resources, there could be more than one cell-edge UE receiving JT-CoMP transmissions on SC $k$ from the same set of coordinated BSs $b, b' \in \mathcal{B}, b \neq b'$, i.e., $|\mathcal{O}^{b,b'}_k| > 1$. A JT-CoMP UE $u \in \mathcal{O}^{b,b'}_k$ in NOMA clusters $C^b_k$ and $C^{b'}_k$ receives its desired signal by applying SIC according to its decoding order $\pi^b_k(u)$ and $\pi^{b'}_k(u)$ in each cluster, respectively. Nevertheless, if a cluster has more than one JT-CoMP UE, the following two conditions are necessary [11]:

- All non-JT-CoMP UEs in NOMA clusters $C^b_k$ and $C^{b'}_k$ should have higher decoding order than UEs receiving JT-CoMP transmissions from BSs $b$ and $b'$. In other words, non-JT-CoMP UEs of both clusters should be able to decode the signals of all JT-CoMP UEs in $\mathcal{O}^{b,b'}_k$ before decoding their desired signals.
- JT-CoMP UEs forming NOMA clusters in multiple cells keep the same decoding order in all cells, regardless of the order of their channel gains.

Accordingly, for a JT-CoMP UE to successfully perform SIC on a SC it is receiving simultaneous signals on, it should maintain the same decoding order in both cells regardless of its channel gains in these cells. Let $\pi^b_k(u)$ and $\pi^{b'}_k(u)$ be the decoding order of JT-CoMP UE $u \in \mathcal{O}^{b,b'}_k$ on SC $k \in \mathcal{K}^b, \mathcal{K}^{b'}$ receiving joint-transmissions from BSs $b, b' \in \mathcal{B}, b \neq b'$, then

$$\pi^b_k(u) = \pi^{b'}_k(u), \forall u \in \mathcal{O}^{b,b'}_k, \forall k \in \mathcal{K}, \forall b, b' \in \mathcal{B}, b \neq b'. \tag{17}$$

## 3. Problem formulation

In this section, the network sum-rate maximization for the resource allocation problem in downlink JT-CoMP-NOMA networks is formulated. In particular, the network sum-rate is obtained as

$$R_{total} = \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} R_{u,k}. \tag{18}$$

Thus, the network sum-of-rates (**SOR**) optimization problem is expressed as,

$$\underline{\textbf{SOR}}: \quad \max_{p^b_{u,k}, \eta^b_{u,k}} R_{total} = \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} R_{u,k}$$

$$\text{s.t.} \quad \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} \eta^b_{u,k} p^b_{u,k} \leq \bar{P}^b, \ \forall b \in \mathcal{B}, \tag{19a}$$

$$R_{u,k} \geq R_{min}, \ \forall k \in \mathcal{K}, \forall u \in \mathcal{U}, \tag{19b}$$

$$\sum_{u \in \mathcal{U}} \eta^b_{u,k} \leq \kappa, \ \forall b \in \mathcal{B}, \forall k \in \mathcal{K}, \tag{19c}$$

$$\sum_{k \in \mathcal{K}} \eta^b_{u,k} \leq \mu, \ \forall b \in \mathcal{B}, \forall u \in \mathcal{U}, \tag{19d}$$

$$|S_u| \leq 1, \ \forall u \in \mathcal{T}, \tag{19e}$$

$$|S_u| \leq \tau, \ \forall u \in \mathcal{E}, \tag{19f}$$

$$p^b_{i,k} \geq p^b_{j,k}, \forall i, j \in \{C^b_k | \pi^b_k(i) < \pi^b_k(j)\}, \quad i \neq j, \forall k \in \mathcal{K}, \forall b \in \mathcal{B}, \tag{19g}$$

$$|\gamma^b_{u,k} - \gamma^b_{u',k}| > \nu^b_k,$$

$$\forall b \in \mathcal{B}, \forall k \in \mathcal{K}, \forall u, u' \in C_k^b, u \neq u', \tag{19h}$$

$$\pi_k^b(u) = \pi_k^{b'}(u),$$

$$\forall u \in \mathcal{O}_k^{b,b'}, \forall k \in \mathcal{K}, \forall b, b' \in \mathcal{B}, b \neq b', \tag{19i}$$

$$0 \leq p_{u,k}^b \leq \eta_{u,k}^b \bar{P}^b,$$

$$\forall b \in \mathcal{B}, \forall k \in \mathcal{K}, \forall u \in \mathcal{U}, \tag{19j}$$

$$\eta_{u,k}^b \in \{0,1\}, \ \forall b \in \mathcal{B}, \forall k \in \mathcal{K}, \forall u \in \mathcal{U}. \tag{19k}$$

Constraint (19a) is the total power budget constraint per BS $b \in \mathcal{B}$, which guarantees that the total allocated power in a cell is limited to $\bar{P}^b$. Constraint (19b) guarantees QoS for all UEs over their assigned SCs. Constraints (19c) and (19d) ensure that each SC can be assigned to at most $\kappa$ UEs, and each UE can occupy at most $\mu$ SCs, respectively. Constraints (19e) and (19f) pose an upper limit on the number of serving BSs for cell-center and cell-edge UEs to 1 and $\tau$ BSs, respectively. Constraint (19g) ensures that power is allocated to UEs according to their decoding order in the ordered list $\pi_k^b, \forall b \in \mathcal{B}, k \in \mathcal{K}$. Constraint (19h) ensures that all users that occupy a SC satisfy the distinct channel gains condition. Constraint (19i) ensures that any cell-edge UE receiving JT-CoMP transmissions from BSs $b$ and $b'$ has the same decoding order in both cells. Constraint (19j), ensures that the power allocated to UE $u$ over SC $k$ in cell $b$ does not exceed the BS's power budget $\bar{P}^b$ when $\eta_{u,k}^b = 1$; otherwise, if $\eta_{u,k}^b = 0$, the constraint reduces to $p_{u,k}^b = 0$. Finally, constraint (19k) ensures that the assignment variable $\eta_{u,k}^b$ is binary.

The formulated problem is a MINLP problem, which is non-convex and combinatorial in nature, due to the existence of the interference terms in the objective function and the binary decision variables for SC assignment. Thus, it is NP-hard and cannot be solved in polynomial-time complexity [28]. Consequently, we decompose the problem into two sub-problems: (1) SC assignment, and (2) power allocation. The SC assignment problem is modeled in Section 4 as a many-to-many two-sided exchange stable matching game with externalities, whereas the power allocation problem is modeled as a DC programming problem in Section 5.

## 4. Sub-carrier assignment

In this section, the many-to-many two-sided exchange stable matching model for SC assignment is discussed. The model satisfies the following constraints of the **SOR** problem, (19b), (19c), (19d), (19e), (19f), (19i) and (19k). Furthermore, a $k$-clique algorithm is developed to form UE groups that can be paired in a NOMA cluster satisfying constraints (19c) and (19h). Then, the SC assignment algorithm is given.

### 4.1. Many-to-many two-sided exchange-stable matching game formulation

We consider the set of UEs $\mathcal{U}$ and the set of SCs $\mathcal{K}$ as two disjoint sets of selfish and rational players aiming to maximize their own utilities. A matching is defined as an assignment of SCs in $\mathcal{K}$ to UEs in $\mathcal{U}$, which can formally be stated in Definition 1 [18,20].

**Definition 1.** Given two disjoint sets of players, $\mathcal{U}$ and $\mathcal{K}$, a many-to-many matching $\Psi$ is a mapping from the set $\mathcal{U} \cup \mathcal{K} \cup \{0\}$ into the set of all subsets of $\mathcal{U} \cup \mathcal{K} \cup \{0\}$ such that for every $u \in \mathcal{U}$ and $k \in \mathcal{K}$, the following conditions are satisfied:

(i) $\Psi(k) \subseteq \mathcal{U}$;
(ii) $\Psi(u) \subseteq \mathcal{K}$;
(iii) $|\Psi(k)| \leq \kappa$;
(iv) $|\Psi(u)| \leq \mu$;
(v) $k \in \Psi(u) \Leftrightarrow u \in \Psi(k)$.

Conditions (i) and (ii) imply that each SC $k$ is matched with a subset of UEs, and each UE $u$ is matched with a subset of SCs, respectively. Conditions (iii) and (iv) state that each SC $k$ can be matched with at most $\kappa$ UEs, and each UE $u$ can be matched with at most $\mu$ SCs, respectively. Condition (v) ensures that a UE $u$ is matched with SC $k$ if and only if SC $k$ is matched with UE $u$.

**Remark 1.** The matching game formulated above is a many-to-many matching game with externalities[6].

Based on (10), it is clear that the data rate of each UE $u$ over each SC $k$ is a function of both intra- and inter-cell interference, i.e., $I_{u,k}$ and $\varphi_{u,k}$, respectively. This coupling between users' utilities due to the interference terms introduces externalities to the matching game. In other words, this means that the utility of each UE depends not only on the SC it is assigned to, but also on the other UEs assigned to the same SC. Additionally, in case of a cell-edge UE $u$, its rate also depends on other UEs assigned to the same SC in neighboring cells that are not necessarily serving it, i.e., not in $\mathcal{S}_u$. Thus, the user pairing over a SC is dependent on the other users in the same cell or other neighboring cells.

**Remark 2.** The utility of a SC $k$ is defined in terms of the sum-rate of the subset of UEs $\mathcal{N} \subset \mathcal{U}$ assigned to it, which can be written as

$$U_k(\mathcal{N}) = \sum_{u \in \mathcal{N}} \log_2\left(1 + \frac{\sum_{b \in \mathcal{S}_u} p_{u,k}^b \gamma_{u,k}^b}{1 + I_{u,k} + \varphi_{u,k}}\right), \forall k \in \mathcal{K}. \tag{20}$$

**Remark 3.** The utility of a UE $u \in \mathcal{U}$ over a SC $k$ is defined as the achievable rate over signals received from all its serving BSs on SC $k$, and is given by

$$U_u(k) = \log_2\left(1 + \frac{\sum_{b \in \mathcal{S}_u} p_{u,k}^b \gamma_{u,k}^b}{1 + I_{u,k} + \varphi_{u,k}}\right), \forall u \in \mathcal{U}, \forall k \in \mathcal{K}. \tag{21}$$

During initialization of the matching process, both UEs and SCs need to set up preference lists in a descending order of preference[7] with respect to their own interests over the opposite set of players, as follows:

i. ***Preference lists of cell-center UEs:*** The preference list $\mathcal{P}_u$ of any cell-center UE $u \in \mathcal{T}$, over the set of SCs can be described as follows. For any two SCs $k, k' \in \mathcal{K}, k \neq k'$, the preference list is given by

$$\mathcal{P}_u: \ k \succ_u k' \Leftrightarrow U_u(k) > U_u(k'), \quad \forall u \in \mathcal{T}, \tag{22}$$

which indicates that a cell-center UE $u$ prefers SC $k$ over $k'$ only if it achieves a higher utility over $k$ than over $k'$.

ii. ***Preference lists of cell-edge UEs:*** The preference list $\mathcal{P}_u$ of any cell-edge UE $u \in \mathcal{E}$, over the set of SCs can be described as follows. For any two subsets of SCs $\mathcal{L}, \mathcal{L}' \subset \mathcal{K}, \mathcal{L} \neq \mathcal{L}'$, the preference list is given by

$$\mathcal{P}_u: \ \mathcal{L} \succ_u \mathcal{L}' \Leftrightarrow U_u(\mathcal{L}) > U_u(\mathcal{L}'), \quad \forall u \in \mathcal{E}, \tag{23}$$

which indicates that a cell-edge UE $u$ prefers the set of SCs $\mathcal{L}$ over $\mathcal{L}'$ only if it achieves a higher utility over $\mathcal{L}$ than over $\mathcal{L}'$. Note that for cell-edge UEs, their preference lists are described over subsets of SCs because they can receive joint transmissions from multiple BSs, possibly over different SCs.

iii. ***Preference lists of SCs:*** The preference list $\mathcal{P}_k$ of any SC $k \in \mathcal{K}$, over the set of UEs can be described as follows. For any two

[6] Externalities are defined as the effects on a player's utility due to the other players' sharing the same commodity [43]. By analogy, a user's rate is affected by the other users sharing the same SC in the same cell or neighboring cells.
[7] $\succ$ denotes a preference relation for UEs and SCs.

subsets of UEs $\mathcal{N}, \mathcal{N}' \subset \mathcal{U}, \mathcal{N} \neq \mathcal{N}'$, the preference list is given by

$$\mathcal{P}_k: \ \mathcal{N} \succ_k \mathcal{N}' \Leftrightarrow U_k(\mathcal{N}) > U_k(\mathcal{N}'), \tag{24}$$

which implies that a SC $k$ prefers the set of UEs $\mathcal{N}$ over $\mathcal{N}'$ only if its utility over $\mathcal{N}$ is higher than that over $\mathcal{N}'$. Note that the preference lists of SCs are described over subsets of UEs due to the presence of externalities in the matching game.

**Definition 2.** A SC $k \in \mathcal{K}$ is deemed **acceptable** to UE $u \in \mathcal{U}$ if $R_{u,k} \geq R_{\min}$.

**Remark 4.** The preference lists of players in the formulated many-to-many matching game are incomplete.[8]

When constructing the preference lists, UEs may declare one or more subsets of SCs as unacceptable, and similarly, SCs may declare one or more subsets of UEs as *unacceptable* [44,45]. Consider the construction of preference lists for cell-edge UEs. As stated above, a cell-edge UE $u \in \mathcal{E}$ will construct its preference list over subsets of SCs because they are able to receive JT-CoMP transmissions from their coordinated BSs. Moreover, recall that in a JT-CoMP system, cell-edge UEs are jointly served by their coordinated BSs using the same SCs. Furthermore, each cell-edge UE will consider subsets of SCs $\mathcal{L} \subset \mathcal{K}$ only from BSs which are within their coverage area, such that $|\mathcal{L}| \leq \tau$ in accordance with constraint (19f). Therefore, any subset of SCs from BSs that are not within the coverage area of $u$ is considered as an unacceptable match for $u$, and thus is not included in its preference list. On the other hand, cell-center UE $u \in \mathcal{T}$ will also construct its preference list over subsets of SCs. However, unlike cell-edge UEs, cell-center UEs will only consider subsets $\mathcal{L} \subset \mathcal{K}$ of cardinality 1, i.e., $|\mathcal{L}| = 1$, as acceptable due to the fact that a cell-center UE does not receive JT from neighboring cells. In addition, only SCs from the BS that the UE is associated with are considered acceptable. Therefore, any subset $\mathcal{L}$ such that $|\mathcal{L}| > 1$ or subset of SCs from other than the associated BS will be considered as an unacceptable match for $u$, and thus will not be included in its preference list. It is important to note that cell-center and cell-edge UEs, consider a subset $\mathcal{L}$ as unacceptable if their utility over $\mathcal{L}$, i.e., $U_u(\mathcal{L})$, does not satisfy (19b).

We now consider the construction of SCs' preference lists. According to (19h), SCs only consider subsets of UEs $\mathcal{N} \subset \mathcal{U}$ who satisfy the distinct channel gains condition. Also, any subset $\mathcal{N}$ that includes a UE which is not within the coverage area of the BS that SC $k$ belongs to is not considered acceptable. Additionally, due to constraint (19c), for $k$ to consider $\mathcal{N}$, the cardinality of $\mathcal{N}$ must be less than or equal to $\kappa$. Finally, for each UE $u \in \mathcal{N}$, the utility of $u$ over $kU_u(k)$ must satisfy the minimum rate constraint. Therefore, any subset of UEs $\mathcal{N}$ that does not satisfy these conditions will be considered as an unacceptable match for $k$, and thus is not included in $k$'s preference list. Hence, for any player, not all subsets of $\mathcal{U} \cup \mathcal{K} \cup \{0\}$ are considered acceptable to match with, and thus, the players' preference lists are *incomplete*. Subsequently, it is possible that a UE $u$ does not get assigned in a given network instance.

**Remark 5.** Unlike traditional matching, the preference lists in the formulated matching game lack the property of substitutability.[9]

Suppose that for a given SC $k \in \mathcal{K}$, its most preferred set of UEs is $\mathcal{N} \subset \mathcal{U}$ which contains $u$ and $u'$, such that $|\gamma^b_{u,k} - \gamma^b_{u',k}| > \nu^b_k$ and $\gamma^b_{u,k} > \gamma^b_{u',k}$. If $u \in \mathcal{N}$, then it is not necessary that $u' \in \mathcal{N} \backslash \{u\}$. Due to the interference terms, the utility $U_{u'}(k)$ of $u'$ over $k$ may have changed after $u$ is removed so that the data rate of $u'$ no longer satisfies the minimum rate requirement, and thus, SC $k$ may not prefer $u'$ any more [20].

Designing matching games becomes significantly more challenging due to the lack of substitutability and the presence of externalities. In fact, the presence of externalities makes it more difficult to determine the existence of stable matchings. Furthermore, if a stable matching exists, it can be computationally difficult to find [47]. Consequently, motivated by the many-to-one "student-housing assignment" [47] and "staff-office assignment" [48] matching problems, we introduce the notion of *two-sided exchange stability (2ES)* to the formulated many-to-many matching game. Different from the traditional definition of stability defined in [49], 2ES is relevant when players are able to compare and exchange their matchings with each other. In order to achieve 2ES, we enable swap matching between the SC assignments of the UEs. Before defining exchange stability, it is convenient to first define *swap matching* as follows [18,20,47]

**Definition 3.** Given a matching $\Psi$ with $k \in \Psi(u)$, $k' \in \Psi(u')$, and $k' \notin \Psi(u)$, $k \notin \Psi(u')$, a swap matching $\Psi^{uk}_{u'k'} = \Psi \backslash \{(u,k),(u',k')\} \cup \{(u,k'),(u',k)\}$ is defined such that $k' \in \Psi^{uk}_{u'k'}(u)$, $k \in \Psi^{uk}_{u'k'}(u')$, and $k \notin \Psi^{uk}_{u'k'}(u)$, $k' \notin \Psi^{uk}_{u'k'}(u')$.

**Remark 6.** One or more of the players participating in a swap could be a "hole".

Note that the players involved in the swap described in Definition 3 are UEs $u$ and $u'$, and SCs $k$ and $k'$, whereas all other matchings remain the same. Moreover, as stated in Remark 6, one of the UEs involved in the swap can be a "hole" [47]—denoted by $H$—representing an open spot in the SC. We elaborate further on the definition of "hole" in Appendix A.

**Remark 7.** Given a matching $\Psi$ and a pair of cell-edge UEs $(u,u')$, receiving JT-CoMP transmissions and matched in $\Psi$, if there exist $\mathcal{L} \in \Psi(u)$, $\mathcal{L}' \in \Psi(u')$, $\mathcal{L}' \notin \Psi(u)$, $\mathcal{L} \notin \Psi(u')$, and $\mathcal{L}, \mathcal{L}' \subset \mathcal{K}, \mathcal{L} \neq \mathcal{L}'$, a swap matching $\Psi^{u\mathcal{L}}_{u'\mathcal{L}'} = \Psi \backslash \{(u,\mathcal{L}),(u',\mathcal{L}')\} \cup \{(u,\mathcal{L}'),(u',\mathcal{L})\}$ is defined such that $\mathcal{L}' \in \Psi^{u\mathcal{L}}_{u'\mathcal{L}'}(u)$, $\mathcal{L} \in \Psi^{u\mathcal{L}}_{u'\mathcal{L}'}(u')$, $\mathcal{L} \notin \Psi^{u\mathcal{L}}_{u'\mathcal{L}'}(u)$, $\mathcal{L}' \notin \Psi^{u\mathcal{L}}_{u'\mathcal{L}'}(u')$.

**Remark 8.** Cell-edge UEs receiving JT-CoMP transmissions can exchange their matchings only with other cell-edge UEs that are also receiving JT-CoMP transmissions.

Remark 7 implies that a swap matching for cell-edge UEs receiving joint transmissions can be accomplished by exchanging the set of SCs that the UEs are being multiplexed on. Moreover, Remark 8 complements the idea by stating that JT-CoMP cell-edge UEs are only able to perform swap operations with other JT-CoMP cell-edge UEs, which is due to the fact that JT-CoMP UEs are matched to subsets of SCs.

**Remark 9.** A cell-center UE can only exchange its SCs with other cell-center UEs or with cell-edge UEs not receiving JT-CoMP transmissions, i.e., non-JT-CoMP. Likewise, a non-JT-CoMP cell-edge UE can only exchange its SCs with other non-JT-CoMP cell-edge UEs or with cell-center UEs.

It is important to note that in order to satisfy 2ES, not only the UEs should be satisfied with the swap, but also the SCs involved should approve the swap. Thus, it is necessary to define the *swap-blocking pair* [18,20].

**Definition 4.** Given a matching $\Psi$ and a pair of UEs $(u,u')$ matched in $\Psi$, if there exist $k \in \Psi(u)$ and $k' \in \Psi(u')$ such that

(1) $\forall i \in \{u,u',k,k'\}, U_i\left(\Psi^{uk}_{u'k'}\right) \geq U_i(\Psi)$ and,

(2) $\exists i \in \{u,u',k,k'\}, U_i\left(\Psi^{uk}_{u'k'}\right) > U_i(\Psi)$,

[8] A preference list of a player is called *complete* if it is based on a strict preference order over *all* the players in the other set of players.

[9] Substitutability can be described as follows [46] Assume that SC $k$ in matching $\Psi$ is assigned to UEs $u'$ and $u''$, such that $\Psi(k) = \{u', u''\}$. Suppose that another UE $u$ exists and it is satisfied to be assigned to SC $k$ in $\Psi$ and join $u'$ and $u''$. Then, UE $u$ must still be satisfied to be assigned to SC $k$ when only $u'$ is assigned to it, i.e., when $\Psi(k) = \{u'\}$.

where $U_i(\Psi)$ represents the utility of the player $i$ under the matching $\Psi$, then swap matching $\Psi^{uk}_{u'k'}$ is approved, and $(u, u')$ is called a swap-blocking pair in $\Psi$.

Observe that Definition 4 implies that if two UEs want to exchange their SCs, then the SCs must approve the swap and hence, ensuring two-sided exchange stability. Condition (1) indicates that the utilities of all players participating in the swap should not decrease after the swap operation. Condition (2) implies that the utility of at least one of the players should increase after the swap operation.

**Remark 10.** For UE pair $(u, u')$ in matching $\Psi$, where both UEs are assigned joint transmissions, if there exist $\mathcal{L}, \mathcal{L}' \subset \mathcal{K}$, $\mathcal{L} \in \Psi(u)$, and $\mathcal{L}' \in \Psi(u')$ such that

(1) $\forall i \in \{u, u', \mathcal{L}, \mathcal{L}'\}, U_i\left(\Psi^{u\mathcal{L}}_{u'\mathcal{L}'}\right) \geq U_i(\Psi)$ and,

(2) $\exists i \in \{u, u', \mathcal{L}, \mathcal{L}'\}, U_i\left(\Psi^{u\mathcal{L}}_{u'\mathcal{L}'}\right) \geq U_i(\Psi)$,

then $(u, u')$ form a swap-blocking pair and swap matching $\Psi^{u\mathcal{L}}_{u'\mathcal{L}'}$ is approved.

Remark 10 follows from Definition 4 and implies that when JT-CoMP UEs want to exchange their subset of SCs, then all SCs $j \in \mathcal{L}$ and $j' \in \mathcal{L}'$ should approve the swap to guarantee 2ES. Subsequently, we can formally define the notion of two-sided exchange stability as follows [47].

**Definition 5.** A matching $\Psi$ is 2ES if and only if a swap-blocking pair does not exist.

According to the above definitions, we can describe the matching process in the considered multi-cell network as follows. The BSs start by checking for prospect swap-blocking pairs by arranging UE pairs. If necessary, BSs will communicate through the back-haul network to exchange information when UE pairs include cell-edge UEs. Thereafter, the BSs will check whether the UEs can benefit by swapping their matchings without affecting the utilities of the SCs involved. Moreover, by searching for swap-blocking pairs and eliminating them by swap matching we are able to reach a stable matching solution according to Definition 5, i.e., 2ES.

### 4.2. Modified k-clique algorithm for UE-group formation

For forming the potential UE groups or subsets that each SC can match with, given that each UE group satisfies the distinct channel gains condition, we introduce the notion of *k-clique* in *undirected graphs* in Section 4.2 and propose a modified Bron–Kerbosch algorithm [29] for finding all possible UE groups. The formation of UE groups necessary for constructing SCs' preference lists $\mathcal{P}_k, \forall k \in \mathcal{K}$ is described in this subsection. Recall that according to constraint (19h), UEs paired in a NOMA cluster should satisfy the distinct channel gains condition to successfully perform SIC. Moreover, the number of UEs multiplexed on a SC is limited to $\kappa$. Consequently, for a SC $k$ in BS $b$, the size of NOMA cluster $\mathcal{C}^b_k$ can range from 1 to $\kappa$,[10] i.e., $1 \leq |\mathcal{C}^b_k| \leq \kappa$.

Let $\mathbb{G}^b_k = (\mathbb{V}^b_k, \mathbb{L}^b_k), \forall k \in \mathcal{K}, \forall b \in \mathcal{B}$ denote a graph comprising a set of vertices $\mathbb{V}^b_k$ representing UEs, the set of undirected edges or links $\mathbb{L}^b_k$ connecting them. A link $l$ between UEs $u$ and $u'$ exists only if (19h) is satisfied, i.e., $\mathbb{L}^b_k = \{l \mid \text{constraint (19h)}\}$. Therefore, graph $\mathbb{G}^b_k$ represents links between UEs that satisfy the distinct channel gains condition on SC $k$ in BS $b$. To extract all possible groups of UEs – that can form a NOMA cluster – from $\mathbb{G}^b_k$, we need to retrieve all *complete sub-graphs* in $\mathbb{G}^b_k$ of size $\kappa$ at most. Before proceeding, it is worthwhile to provide the following graph theory definitions and remarks [29,50].

[10] Note that a SC can be assigned to no UE. This depends on the preference lists of the players and the matching process, as will be discussed in Section 4.3.

**Definition 6.** A sub-graph $\mathbb{S}^b_k$ composed of vertices of $\mathbb{G}^b_k$ is complete only if each vertex in $\mathbb{S}^b_k$ has a link with every other vertex in $\mathbb{S}^b_k$.

**Definition 7.** A maximal complete sub-graph or clique is a complete sub-graph that is not contained in any other complete sub-graph.

**Remark 11.** A clique with 'k' vertices is usually referred to as a k-clique.

**Algorithm 1:** k-clique-based UE-Group Formation Algorithm (K-UGFA)

**Step 1:** ***Initialization***

1: Obtain CSI for all UEs in the network;
2: Initialize the set of cliques found, $\mathbb{S}^b_k = \emptyset, \forall k \in \mathcal{K}, \forall b \in \mathcal{B}$;

**Step 2:** ***Clique Formation Process***

3: For each $k \in \mathcal{K}, b \in \mathcal{B}$ construct $\mathbb{G}^b_k = (\mathbb{V}^b_k, \mathbb{L}^b_k)$;
4: **for** $\forall v \in \mathbb{V}^b_k$ **do**
5: $v$ is added to the set $\mathbb{S}^b_k$ as a 1-clique, $\mathbb{S}^b_k = \mathbb{S}^b_k \cup \{(v)\}$;
6: Call **Algorithm 2**: $\mathbb{S}^b_k = \mathbb{S}^b_k \cup$ **RICSF**$\left((v), q, \mathbb{G}^b_k, \kappa\right)$, where $q$ is the size of clique $(v, v')$, i.e., $q = 1$;
7: **end for**

**Output:** The set of possible cliques $\mathbb{S}^b_k, \forall k \in \mathcal{K}, \forall b \in \mathcal{B}$;

**Algorithm 2:** Recursive Increase Clique Size Function (RICSF)

**Input:** $Q, q, \mathbb{G}^b_k, \kappa$
1: **if** $q = \kappa$ **then**
2: exit;
3: **else**
4: Retrieve all vertices $\hat{v}$ that could be added to $Q$ to form a new clique $Q^*$ of size $(q + 1)$;
5: For each $\hat{v}$, $Q^* = Q + \hat{v}$;
6: **RICSF**$(Q^*, q + 1, \mathbb{G}^b_k, \kappa)$;
7: **end if**

**Output:** A list which includes all cliques – of size less than or equal to $\kappa$ – originated from clique $Q$;

**Remark 12.** Any vertex that is not connected to any other vertex, i.e., isolated, is considered as a clique of size 1, i.e., 1-clique.

From Definition 6, it becomes evident that potential UE groups should form complete sub-graphs to ensure that (19h) is fulfilled between all UEs in a group. Accordingly, our objective is to list all possible cliques in graph $\mathbb{G}^b_k$, i.e., from 1-clique to $\kappa$-clique, in order to find all possible UE groups that could form a NOMA cluster on SC $k$ in BS $b$. Inspired by Bron–Kerbosch (BK) algorithm for finding all maximal cliques of an undirected graph [29], we propose a modified clique algorithm to enumerate all possible k-cliques in graph $\mathbb{G}^b_k$ that have a maximum of $\kappa$ vertices in order to satisfy constraint (19c). Note that the original BK algorithm is implemented to find all maximal complete sub-graphs; whereas, our objective is to find all cliques that have a maximum of $\kappa$ vertices – regardless of whether the clique is a maximal clique or not – in order to satisfy constraint (19c). **Algorithm 1** presents the details of the proposed k-clique-based UE-Group Formation Algorithm (**K-UGFA**).

**Algorithm 1** outlines the pseudocode of the proposed k-clique-based UE-Group Formation Algorithm (**K-UGFA**). The proposed **K-UGFA** algorithm starts by adding all vertices in graph $\mathbb{G}^b_k$ as a 1-clique to the set of found cliques $\mathbb{S}^b_k$. Thereafter, each 1-clique needs to be increased in size by adding vertices to it while making sure that the resulting sub-graph is still complete. Each resulting sub-graph is increased in size until we reach a $\kappa$-clique. This process is efficiently achieved by using the Recursive Increase Clique Size Function (**RICSF**) presented in **Algorithm 2**. The **RICSF** function takes the following input arguments: the clique to be increased $Q$, its size $q$, the original graph $\mathbb{G}^b_k$ and the SC quota $\kappa$. Each call of the **RICSF** function returns a clique $Q^*$ of size $(q + 1)$. Subsequently the function is recursively called

to increase the size of $Q^*$, and so on. This process continues until the size of the returned clique is equal to $\kappa$, i.e., $q = \kappa$.

The computational complexity of **RICSF** and K-**UGFA** are presented in Appendices B and C, respectively.

**Algorithm 3:** Initialization UE-SC Matching Algorithm (IUSMA)

**Step 1:** ***Initialization***

Initialize UE groups using **K-UGFA**;
Construct UEs preference lists $\mathcal{P}_u, \forall u \in \mathcal{U}$ and SCs preference lists $\mathcal{P}_k, \forall k \in \mathcal{K}$;
Set of unmatched UEs $\mathcal{Z} = \{u | \mathcal{P}_u \neq \emptyset, |\Psi(u)| < \mu, \forall u \in \mathcal{U}\}$, set of JT-CoMP UEs $\mathcal{M} = \emptyset$, and matching state $\Psi = \emptyset$;

**Step 2:** ***Matching Process***

**while** $\mathcal{Z} \neq \emptyset$ **do**
  **for** every unmatched UE $u \in \mathcal{Z}$ **do**
    **if** $u \in \mathcal{T}$ **then**
      UE $u$ proposes to its most preferred SC that has never rejected it before.;
    **else if** $u \in \mathcal{E}$ and $\Psi(u) = \emptyset$ **then**
      UE $u$ proposes to its most preferred SC $\mathcal{L} \in \mathcal{P}_u$ that has never rejected it before;
      **if** $|\mathcal{L}| = \tau$ **then** $\mathcal{M} = \mathcal{M} \cup \{u\}$; **end if**
    **else if** $u \in \mathcal{E}$ and $\Psi(u) \neq \emptyset$ **then**
      **if** $u \notin \mathcal{M}$ **then** UE $u$ proposes to its most preferred SC $\mathcal{L} \in \mathcal{P}_u$, $|\mathcal{L}| = 1$, that has never rejected it before;
      **else** UE $u$ proposes to its most preferred SC $\mathcal{L} \in \mathcal{P}_u$, $|\mathcal{L}| = \tau$, that has never rejected it before; **end if**
    **end if**
  **end for**
  **for** every SC $k \in \mathcal{K}$ **do**
    **if** $k$ is matched to more than 1 UE group **then**
      SC $k$ keeps the most preferred UE group and rejects the others;
      **if** $k$ is rejecting a JT-CoMP cell-edge UE $u'$ **then** $k'$ must also reject $u'$, where $k'$ is defined such that $k \in \mathcal{K}^b$, $k' \in \mathcal{K}^{b'}, k = k'$, and $b, b' \in \mathcal{B}, b \neq b'$; **end if**
    **else**
      SC $k$ keeps its matching;
    **end if**
  **end for**
  Update the matching state $\Psi$ and the set of unmatched UEs $\mathcal{Z}$;
**end while**

**Output:** Initial matching $\Psi$; Set of JT-CoMP UEs $\mathcal{M}$;

### 4.3. Proposed sub-carrier assignment algorithm

The proposed Initialization UE-SC Matching Algorithm (**IUSMA**) is presented in **Algorithm 3**. It constructs the initial matching between UEs and SCs based on many-to-many "firm-worker assignment" matching problem in [51]. After initial matching, we enable swap operations by implementing and applying the Swap Matching Algorithm (**SMA**) to handle the lack of substitutability and account for externalities, ultimately yielding a 2ES matching. In **Algorithm 3**, UE groups are obtained using **K-UGFA**. Subsequently, UEs and SCs construct their preference lists $\mathcal{P}_u$ and $\mathcal{P}_k$, respectively. The set of unmatched UEs, with non-empty preference lists, and the set of CoMP UEs are denoted by $\mathcal{Z}$ and $\mathcal{M}$, respectively. During the matching process, each unmatched UE proposes to its most preferred SC, then each SC accepts the most preferred group of UEs and reject the remaining ones. This process continues until all UEs are either matched or have proposed to all SCs in their preference lists and been rejected, i.e., $\mathcal{P}_u = \emptyset$.

The proposed **SMA** is presented in **Algorithm 4**. After obtaining the initial UE-SC matching from **IUSMA**, each UE $u$ searches for another UE $u'$ or a hole $H_{u'}$ to check whether it forms a swap-blocking pair $(u, u')$ or $(u, H_{u'})$, respectively. If a swap-blocking pair is found, the swap operation is approved and swap matching is executed. The process continues until no UEs can form new swap-blocking pairs, and hence a 2ES matching is obtained.

**Algorithm 4:** Swap Matching Algorithm (SMA)

**Step 1:** ***Initialization***

Initial matching $\Psi$ and set of JT-CoMP UEs $\mathcal{M}$ using **IUSMA**;

**Step 2:** ***Swap Matching Process***

**repeat**
  For each UE $u$, it searches for another UE $u'$ (or a hole $H_{u'}$) to check whether it forms a swap-blocking pair;
  **if** $(u, u')$ or $(u, H_{u'})$ forms a swap-blocking pair in $\Psi$ **then**
    Approve swap operation and update the current matching state (considering Definition 3 and Remarks 6, 7, 8, 9) as follows: $\Psi = \Psi^{u\mathcal{L}}_{u'\mathcal{L}'}$ or $\Psi = \Psi^{u\mathcal{L}}_{H_{u'}\mathcal{L}'}$ or $\Psi = \Psi^{uH_{\mathcal{L}}}_{H_{u'}\mathcal{L}'}$, where $\mathcal{L} \subset \mathcal{K}$ and $|\mathcal{L}| = 1$, $u, u' \notin \mathcal{M}$, or $|\mathcal{L}| = \tau$, $u, u' \in \mathcal{M}$;
  **else**
    Keep the current matching state;
  **end if**
**until** there is no swap-blocking pair

**Output:** 2ES matching $\Psi^* = \Psi$;

The algorithmic properties of the proposed **IUSMA** and **SMA** in terms of stability, convergence, complexity and optimality are presented in Appendices D–H, respectively.

## 5. Power allocation

In this section, we consider the power allocation for UEs in a multi-cell network. Suppose that the SC assignment for all users is obtained, i.e., $\eta^b_{u,k}$ for all UEs over all SCs for all BSs are set, and sets $\mathcal{S}_u, \forall u \in \mathcal{U}$ and $\mathcal{C}^b_k, \forall k \in \mathcal{K}, \forall b \in \mathcal{B}$ are determined. Subsequently, the **SOR** problem reduces to the following power allocation (**PA**) problem,

$$\underline{\textbf{PA}}: \quad \max_{p^b_{u,k}} \quad R_{total} = \sum_{u\in\mathcal{U}} \sum_{k\in\mathcal{K}} R_{u,k}$$

$$\text{s.t.} \quad \text{(19a), (19b), (19g), and (19j).} \tag{25}$$

The minimum rate constraint in (19b) can be transformed into the following linear constraint

$$\sum_{b\in\mathcal{S}_u} \eta^b_{u,k} p^b_{u,k} \gamma^b_{u,k} - \left(2^{R_{min}} - 1\right)\left(1 + I_{u,k} + \varphi_{u,k}\right) \geq 0, \tag{26}$$

which applies for $\forall k \in \mathcal{K}, \forall u \in \mathcal{U}$. Consequently, problem **PA** can be rewritten as follows

$$\max_{p_{u,k}} \quad \sum_{u\in\mathcal{U}} \sum_{k\in\mathcal{K}} R_{u,k} \tag{27}$$

$$\text{s.t.} \quad \text{(19a), (19g), (19j), and (26).}$$

Although all constraints of the current problem are linear, problem (27) is still non-convex due to the non-convexity of the objective function, i.e., network sum-rate [26]. Denote the objective function (25) in problem **PA** by $V(\mathbf{p})$ and rewrite it as a difference of two convex functions

$$V(\mathbf{p}) = f(\mathbf{p}) - g(\mathbf{p}), \tag{28}$$

where $f(\mathbf{p})$ and $g(\mathbf{p})$ are, respectively, defined as follows

$$f(\mathbf{p}) = \sum_{u\in\mathcal{U}} \sum_{k\in\mathcal{K}} \log_2\left(1 + I_{u,k} + \varphi_{u,k} + \sum_{b\in\mathcal{S}_u} \eta^b_{u,k} p^b_{u,k} \gamma^b_{u,k}\right), \tag{29}$$

and

$$g(\mathbf{p}) = \sum_{u\in\mathcal{U}} \sum_{k\in\mathcal{K}} \log_2\left(1 + I_{u,k} + \varphi_{u,k}\right). \tag{30}$$

Both $f(\mathbf{p})$ and $g(\mathbf{p})$ are concave in $\mathbf{p}$ [30], and thus the objective function is transformed into a DC programming problem. The function $g(\mathbf{p})$ can be approximated by its first-order Taylor expansion at $\mathbf{p}^{(t)}$, i.e., $g(\mathbf{p}^{(t)}) + \nabla g^T(\mathbf{p}^{(t)})(\mathbf{p} - \mathbf{p}^{(t)})$ in each iteration $t$ [52,53], where $\nabla g^T(\mathbf{p}^{(t)})$ denotes the gradient of $g(\mathbf{p})$ at $\mathbf{p}^{(t)}$, where $\mathbf{p}^{(t)}$ is a vector of

feasible powers at iteration $t$, and $(\cdot)^T$ denotes the transpose operator. Therefore, problem **PA** is further transformed to

$$\max_{\mathbf{p}} \; V(\mathbf{p}) = f(\mathbf{p}) - g(\mathbf{p}^{(t)}) - \nabla g^T(\mathbf{p}^{(t)})(\mathbf{p} - \mathbf{p}^{(t)}) \tag{31}$$

s.t. (19a), (19g), (19j), and (26),

which is also a concave maximization problem [54].

It is noteworthy that function $g(\mathbf{p})$'s sensitivity to change in the variable $\mathbf{p}$ is relatively low [52]. This makes it possible to efficiently approximate $g(\mathbf{p})$ by its first order Taylor expansion at a large neighborhood of $\mathbf{p}^{(t)}$.

To evaluate the gradient of $g(\mathbf{p})$ at each $\mathbf{p}$ for all UEs $u \in \mathcal{U}$, we decompose the summation over all UEs in (30) into two summations, the first is over cell-center UEs whereas the second is over cell-edge UEs as follows[11]

$$g(\mathbf{p}) = \sum_{u\in\mathcal{T}}\sum_{k\in\mathcal{K}} \log_2\left(1 + I_{u,k}\right) + \sum_{u\in\mathcal{E}}\sum_{k\in\mathcal{K}} \log_2\left(1 + I_{u,k} + \varphi_{u,k}\right). \tag{32}$$

For notational simplicity, define the following:

$$g_1(\mathbf{p}) = \sum_{u\in\mathcal{T}}\sum_{k\in\mathcal{K}} \log_2\left(1 + I_{u,k}\right),$$

and

$$g_2(\mathbf{p}) = \sum_{u\in\mathcal{E}}\sum_{k\in\mathcal{K}} \log_2\left(1 + I_{u,k} + \varphi_{u,k}\right).$$

Subsequently, the gradient of $g_1(\mathbf{p})$ at each $\mathbf{p}$ for cell-center UEs is given by

$$\frac{\partial g_1(\mathbf{p})}{\partial p^m_{i,j}} = \sum_{u\in\mathcal{T}}\sum_{k\in\mathcal{K}} \frac{\alpha^b_{u,k}(i,j,m)}{\ln 2(1 + I_{u,k})}, \quad \forall i \in \mathcal{U}, \forall j \in \mathcal{K}, \forall m \in \mathcal{B}, \forall b \in \mathcal{S}_u, \tag{33}$$

where $\alpha^b_{u,k}(i,j,m)$ is given by

$$\alpha^b_{u,k}(i,j,m) = \begin{cases} \eta^b_{u,k}\gamma^b_{u,k}, & \text{if } i \neq u, j = k, m = b, \\ & \text{for } i \in \{C^b_k | \pi^b_k(i) > \pi^b_k(u)\}, \\ 0, & \text{otherwise.} \end{cases} \tag{34}$$

On the other hand, the gradient of $g_2(\mathbf{p})$ at each $\mathbf{p}$ for cell-edge UEs is given by

$$\frac{\partial g_2(\mathbf{p})}{\partial p^m_{i,j}} = \sum_{u\in\mathcal{E}}\sum_{k\in\mathcal{K}} \frac{\beta^b_{u,k}(i,j,m)}{\ln 2(1 + I_{u,k} + \varphi_{u,k})}, \quad \forall i \in \mathcal{U}, \forall j \in \mathcal{K}, \forall m \in \mathcal{B}, \forall b \in \mathcal{S}_u, \tag{35}$$

where $\beta^b_{u,k}(i,j,m)$ is given by

$$\beta^b_{u,k}(i,j,m) = \begin{cases} \eta^m_{u,k}\gamma^m_{u,k}, & \text{if } i \neq u, j = k, m \in \mathcal{S}_u, \\ & \text{for } i \in \{C^m_k | \pi^m_k(i) > \pi^m_k(u)\}, \\ (1 - \eta^m_{u,k})\gamma^m_{u,k}, & \text{if } i \neq u, j = k, m \notin \mathcal{S}_u, \\ 0, & \text{otherwise.} \end{cases} \tag{36}$$

Thus, in general, the gradient of $g(\mathbf{p})$ at each $\mathbf{p}$ is given by

$$\nabla g(\mathbf{p}) = \underbrace{\sum_{u\in\mathcal{T}}\sum_{k\in\mathcal{K}} \frac{\alpha^b_{u,k}(i,j,m)}{\ln 2(1 + I_{u,k})}}_{\nabla g_1(\mathbf{p})} + \underbrace{\sum_{u\in\mathcal{E}}\sum_{k\in\mathcal{K}} \frac{\beta^b_{u,k}(i,j,m)}{\ln 2(1 + I_{u,k} + \varphi_{u,k})}}_{\nabla g_2(\mathbf{p})}, \quad \forall i \in \mathcal{U}, \forall j \in \mathcal{K}, \forall m \in \mathcal{B}, \forall b \in \mathcal{S}_u. \tag{37}$$

By rewriting problem **PA** in its DC equivalent form with linear constraints, as presented in (31), we are able to locate the corresponding optimal solutions by applying the Frank–Wolfe (FW) algorithm [31]. Initialized at a feasible solution $\mathbf{p}^{(0)}$, the FW algorithm generates a sequence $\{\mathbf{p}^{(t)}\}$ of improved feasible solutions, where at the $t$th iteration, $\mathbf{p}^{(t+1)}$ is generated as the optimal solution [52] of the convex program (31). Problem (31) is a standard convex optimization problem which can be efficiently solved by standard software packages such as CVX [55]. The proposed power allocation algorithm is outlined in **Algorithm 5**.

**Algorithm 5:** FW-based DC Programming Power Allocation Algorithm (FW-DCPA)

**Step 1: *Initialization***
1: Set iteration index $t = 0$ and error tolerance $\epsilon \ll 1$.
2: Initialize $\mathbf{p}^{(0)}$ by allocating equal transmit powers to UEs and calculate $V(\mathbf{p}^{(0)}) = f(\mathbf{p}^{(0)}) - g(\mathbf{p}^{(0)})$, or $R_{total}(\mathbf{p}^{(0)})$;
**Step 2: *Power Allocation***
3: **repeat**
4: Solve problem (31) to obtain the optimal solution $\mathbf{p}^*$;
5: Set $\mathbf{p}^{(t+1)} = \mathbf{p}^*$ and calculate $V(\mathbf{p}^{(t+1)}) = f(\mathbf{p}^{(t+1)}) - g(\mathbf{p}^{(t+1)})$, or $R_{total}(\mathbf{p}^{(t+1)})$;
6: Increment iteration index, $t = t + 1$;
7: **until** $|R_{total}(\mathbf{p}^{(t+1)}) - R_{total}(\mathbf{p}^{(t)})| \leq \epsilon$, or $|\mathbf{p}^{(t+1)} - \mathbf{p}^{(t)}| \leq \epsilon$
**Output:** $\mathbf{p}^{opt} = \mathbf{p}^{(t+1)}$;

In **Algorithm 5**, the solution $\mathbf{p}^{(t+1)}$ is generated as the optimal solution of problem (31) using solution of the previous iteration, i.e., $\mathbf{p}^{(t)}$. As problem (31) is derived from (28), then the solution $\mathbf{p}^{(t)}$ at any iteration $t$ is also feasible to (28). It follows that, $f(\mathbf{p}^{(t+1)}) - g(\mathbf{p}^{(t+1)}) \geq f(\mathbf{p}^{(t)}) - g(\mathbf{p}^{(t)}) - \nabla g^T(\mathbf{p}^{(t)})(\mathbf{p}^{(t+1)} - \mathbf{p}^{(t)}) \geq f(\mathbf{p}^{(t)}) - g(\mathbf{p}^{(t)})$ [26,52]. Consequently, it is clear that $(t+1)$ iteration the next solution $\mathbf{p}^{(t+1)}$ is always better than the $t$th iteration solution $\mathbf{p}^{(t)}$, which implies that the objective function $V(\mathbf{p})$ improves in each iteration.

Furthermore, the set of constraints in (31) is compact and thus, by Cauchy theorem, the sequence of generated $\mathbf{p}^{(t)}$ always converges [52]. Therefore, **Algorithm 5** terminates in a finite number of iterations at $|V(\mathbf{p}^{(t+1)}) - V(\mathbf{p}^{(t)})| \leq \epsilon$ indicating no solution improvement, or at $|\mathbf{p}^{(t+1)} - \mathbf{p}^{(t)}| \leq \epsilon$ when there is no solution movement.

The computational complexity of **FW-DCPA** is presented in Appendix I.

## 6. Joint sub-carrier assignment and power allocation algorithm

In this section, we discuss the general procedure of the proposed joint SC assignment and power allocation scheme (**JSAPA**), which is outlined in **Algorithm 6**.

In the initialization step, all BSs collect UEs' CSI information, categorize UEs into cell-center and cell-edge UEs, set the values of $\kappa, \mu, \tau$ and $\nu^b_k$, allocate equal transmit power to all UEs, and then form all possible UE groups using **Algorithm 1**. Subsequently, each BS constructs the preference lists for all SCs and UEs in its range. This is followed by executing **Algorithm 3** to find an initial matching between UEs and SCs. Then, transmit power for all UEs is updated using **Algorithm 5**. Thereafter, swap operations are executed to eliminate swap blocking pairs while updating the transmit power for all UEs after each approved swap operation. Once all swap blocking pairs are eliminated, the scheme terminates and outputs a 2ES matching between UEs and SCs with the updated transmit power.

It is worthwhile to note that **JSAPA** can be readily used for resource allocation in conventional NOMA (C-NOMA) networks by setting $\tau = 1$ to disable joint-transmissions for cell-edge UEs. Furthermore, by additionally setting $\kappa = 1$ the problem is transformed into a resource allocation problem for OFDMA networks. Hence, the proposed scheme can be used as a generic resource allocation scheme for JT-CoMP-NOMA, C-NOMA and OFDMA networks.

The computational complexity of **JSAPA** is presented in J. In addition, the convergence analysis of **JSAPA** is presented in K.

[11] Recall that for cell-center users, $\varphi_{u,k} = 0$.

**Algorithm 6:** Joint Sub-carrier and Power Allocation Algorithm (JSAPA)

**Step 1:** ***Initialization***

1: The BSs obtains CSI of all UEs in the network;
2: The BSs categorize UEs into cell-center and cell-edge according to (1);
3: Set the values of $\kappa, \mu, \tau$ and $v_k^b$;
4: Equally allocate the BS power to UEs;
5: Each BS forms all possible UE groups that can form possible NOMA clusters using **K-UGFA** (**Algorithm 1**);

**Step 2:** ***Construct the preference lists***

6: Each BS constructs the preference lists for cell-center UEs (22), cell-edge UEs (23), and its SCs (24);

**Step 3:**

7: Construct initial matching between UEs and SCs using **IUSMA** (**Algorithm 3**);
8: Update the power allocation vector **p** using **FW-DCPA** (**Algorithm 5**) according to the initial matching;
9: **repeat**
10: Enable swap operations using **SMA** (**Algorithm 4**) to eliminate swap blocking pairs;
11: Update the power allocation vector **p** using **FW-DCPA** (**Algorithm 5**) after each approved swap;
12: **until** all swap blocking pairs are eliminated

**Output:** SC assignment $\eta_{u,k}^b, \forall u \in \mathcal{U}, \forall k \in \mathcal{K}, \forall b \in \mathcal{B}$, and power allocation vector **p**;

**Table 2**
Simulation parameters.

| Simulation parameter | Value |
|---|---|
| Cell radius | 215 m |
| Total transmit power of BSs ($\bar{P}^b$) | 43 dBm |
| Path loss exponent ($\alpha$) | 3, urban area |
| Threshold $\gamma^{thr}$ | 20.628 dBm |
| Noise spectral density ($\sigma^2$) | −50 dBm/Hz |
| Minimum rate ($R_{min}$) | 0.3 bps/Hz |
| Error tolerance ($\epsilon$) | $10^{-3}$ |
| Number of coordinated BSs ($\tau$) | 2 |

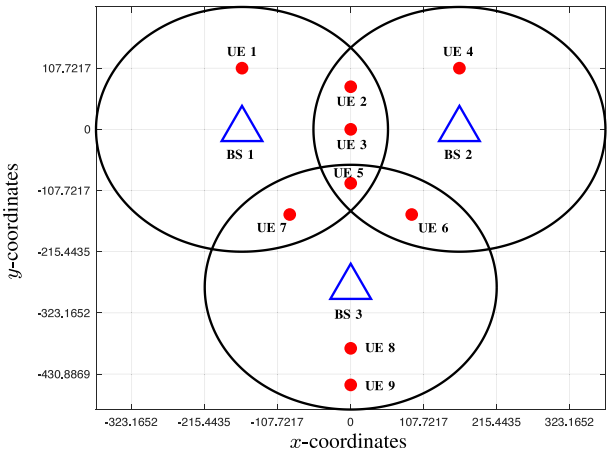


**Fig. 1.** Network topology. BS locations are labeled by triangles at the center of the cells, while UEs locations are labeled by circles.

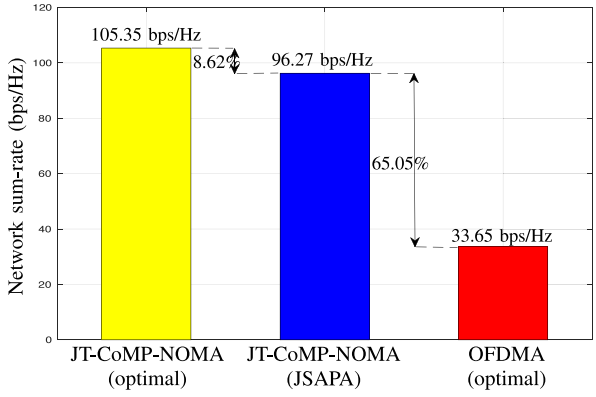


**Fig. 2.** Network sum-rate comparison for JT-CoMP-NOMA (optimal), JT-CoMP-NOMA (JSAPA) and OFDMA (optimal).

## 7. Simulation results

In this section, we evaluate the performance of the proposed resource allocation algorithm in JT-CoMP-NOMA networks and compare it against a baseline solution computed by a commercial optimization package [32], in addition to other schemes like C-NOMA and OFDMA. To the best of the authors' knowledge, no prior work has considered similar problem formulation and constraints to our work. Therefore, comparisons to other schemes from the literature would not be possible.

The main parameters used in our simulations are shown in Table 2 [7,18,56,57]. The cell-edge threshold is set to $\gamma^{thr} = 20.628$ dBm, which is the average SNR measured at the edge of the neighboring cell that falls within the coverage area of a given cell. The user pairing threshold $v_k^b$ is set to the standard deviation of the UEs channel gains, i.e., $v_k^b = \text{std}(\mathbf{H}_k^b), \forall k \in \mathcal{K}, b \in \mathcal{B}$, where $\mathbf{H}_k^b = [h_{1,k}^b, \ldots, h_{u,k}^b]$. In order to evaluate the performance of the proposed algorithm in OFDMA networks, we set $\kappa = 1$ and $\tau = 1$.

To evaluate the performance of the proposed algorithm, we setup a small network as shown in Fig. 1, and compare the results against a baseline solution computed by a polyhedral branch-and-cut approach [58]. This approach is implemented in a commercial package called "BARON", which is available through the General Algebraic Modeling System (GAMS)[12] [59]. The network consists of a total of 9 UEs distributed among the cells, where each cell has 5 UEs within its coverage area. The spectrum allocated to each BS is divided into $K = 3$ SCs, whereas $\mu$ and $\kappa$ are set as follows: $\mu = 2$ and $\kappa = 2$.

[12] For details on the benchmark implementation we refer the reader to [32].

Due to the complexity of the **SOR** problem and the considerable amount of time needed to arrive at the optimal solution, we provide comparison of the network sum-rate by averaging the user channel gains over 1000 network instances. Fig. 2 shows a comparison of the network sum-rate for JT-CoMP-NOMA (optimal), JT-CoMP-NOMA (JSAPA) and OFDMA (optimal). Results demonstrate that the proposed **JSAPA** scheme achieves 91.38% of a baseline solution. In addition, **JSAPA** achieves an increase in sum-rate equivalent to 65.05% of the sum-rate achieved by OFDMA (optimal). The **JSAPA** maintains a balanced trade-off between convergence speed and stability in one side, and quality of the solution on the other side.

For the next set of results, we set up a multi-cell network composed of 3 circular cells, where the BSs are placed at the center of the cells. The frequency spectrum allocated to each BS is divided into $K = 20$ SCs. The results provided are evaluated by averaging over 1000 independent network instances, where **JSAPA** scheme is applied to optimize resource allocation for JT-CoMP-NOMA, C-NOMA and OFDMA. In each simulated network instance, UEs are randomly scattered within the cells and the channel coefficients are randomly generated for each network instance.

In the first set of experiments, we allow each UE in the network to be allocated at most three SCs, i.e., $\mu = 3$. For the JT-CoMP-NOMA and C-NOMA schemes, we allow a maximum of 2 UEs to be multiplexed per SC, i.e., $\kappa = 2$, whereas for OFDMA $\kappa = 1$. Moreover, for C-NOMA and OFDMA $\tau = 1$. Fig. 3 illustrates the average network sum-rate, the average data rate achieved per cell-center UE and the average data rate achieved per cell-edge UE versus the number of UEs in the network. It is

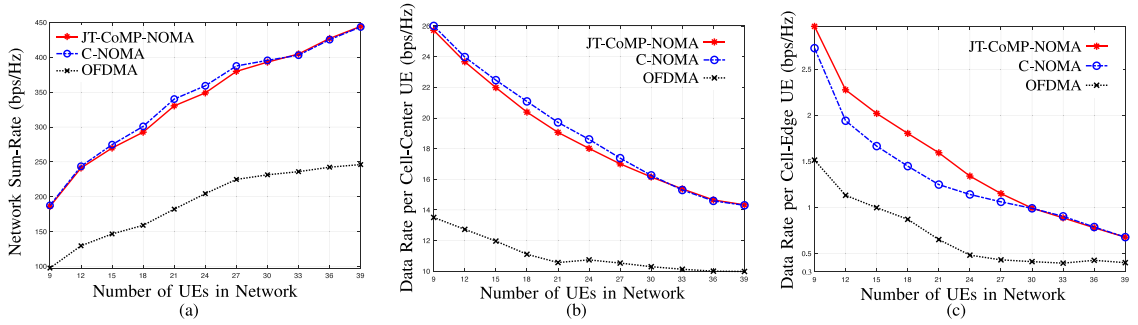


**Fig. 3.** Data rates in JT-CoMP-NOMA ($\mu = 3, \kappa = 2, \tau = 2$), C-NOMA ($\mu = 3, \kappa = 2, \tau = 1$) and OFDMA ($\mu = 3, \kappa = 1, \tau = 1$) networks vs. the total number of UEs in the network.

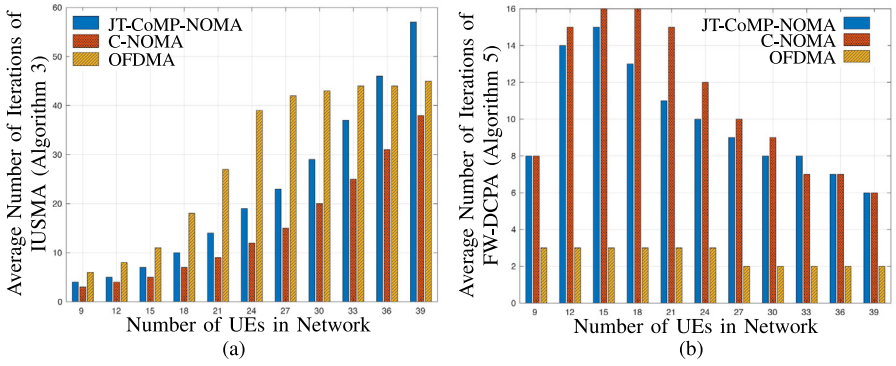


**Fig. 4.** Average Number of Iterations for JT-CoMP-NOMA ($\mu = 3, \kappa = 2, \tau = 2$), C-NOMA ($\mu = 3, \kappa = 2, \tau = 1$) and OFDMA ($\mu = 3, \kappa = 1, \tau = 1$) schemes.

clear from Fig. 3(a) that the network sum-rate increases as the number of UEs in the network increases. Moreover, JT-CoMP-NOMA and C-NOMA schemes show significant improvement in the network sum-rate over OFDMA as more UEs are able to simultaneously utilize the same SC. However, it is noticed that by enabling JT-CoMP in NOMA, the network sum-rate slightly decreases as compared to C-NOMA. A similar observation can be noted in Fig. 3(b), where cell-center UEs in JT-CoMP-NOMA experience a slight decrease in their achievable rates as compared to cell-center UEs in C-NOMA. Nevertheless, as shown in Fig. 3(c), the improvement in data rates of cell-edge UEs in JT-CoMP-NOMA ranges from 0.1 – 27.7% in comparison to C-NOMA. This improvement is especially noticeable at lower numbers of UEs in the network. As the number of UEs in the network increases, this difference becomes minimal. Therefore, the trade-off in enabling JT-CoMP is that we are able to improve cell-edge UEs' data rates and mitigate the inter-cell interference for such users.

Fig. 4 illustrates the average number of iterations of the proposed algorithms versus the number of UEs in the network for JT-CoMP-NOMA, C-NOMA and OFDMA. Fig. 4(a) shows the average number of iterations of the many-to-many matching algorithm outlined in **IUSMA** (**Algorithm 3**). It is observed that the number of iterations increases as the number of UEs in the network increases. This is because larger number of players participating in the matching game results in additional searching dimensions in the potential matching solutions. Moreover, some UEs might need – in the worst case scenario – to propose to all SCs in their preference lists in order to find whether a possible match is available. The JT-CoMP-NOMA has the largest number of iterations due to the larger number of possibilities that cell-edge UEs need to consider, i.e., JT-CoMP and non-JT-CoMP combinations of SCs, as discussed in Section 4.1. An interesting observation is that OFDMA requires more iterations than C-NOMA to converge. This is due to the fact that UEs in OFDMA cannot share spectrum resources simultaneously, which forces higher competition among the UEs in the network. Therefore, it takes more iterations to match a UE with its best possible SC. However, this competition saturates as the number of UEs becomes larger than the number of SCs.

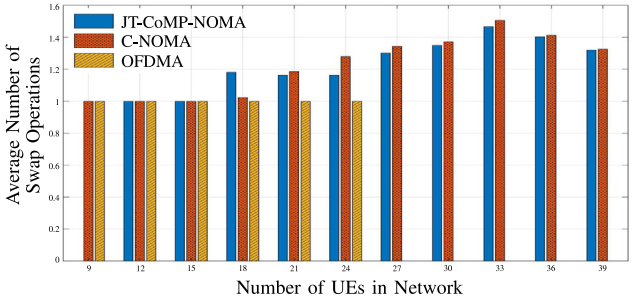


**Fig. 5.** Average Number of Swap Operations for JT-CoMP-NOMA ($\mu = 3, \kappa = 2, \tau = 2$), C-NOMA ($\mu = 3, \kappa = 2, \tau = 1$) and OFDMA ($\mu = 3, \kappa = 1, \tau = 1$) schemes.

Fig. 4(b) shows the average number of iterations of the power allocation algorithm presented in **FW-DCPA** (**Algorithm 5**). It can be seen that the number of iterations decreases as the number of UEs in the network increases. This can be justified as follows, since the spectrum and power resources are fixed, the power feasible region decreases as the number of UEs increases, and thus fewer iterations are required for the algorithm to converge.

Fig. 5 presents the average number of swap operations required to account for the externalities. It can be seen that JT-CoMP-NOMA and C-NOMA require on average a single swap operation for small number of UEs in the network. As the number of UEs increases, the average number of swaps marginally increases up to a certain level where it starts to decrease again. On the other hand, OFDMA requires a single swap operation for small number of UEs in the network, while no swaps are required for large number of UEs. This pattern in the results occurs because, when the number of UEs increases, there are more possibilities that UEs could form swap blocking pairs and exchange their matched SCs. However, further increase in the number of UEs reduces the possibility of performing swaps because resources become

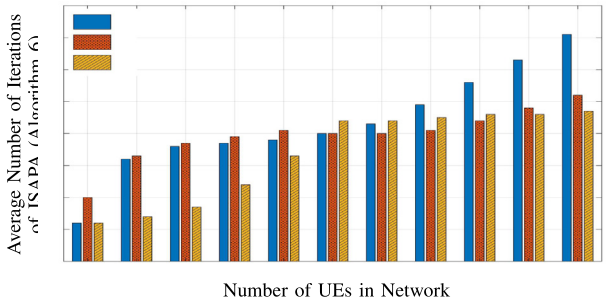


**Fig. 6.** Average total number of iterations of **JSAPA**.

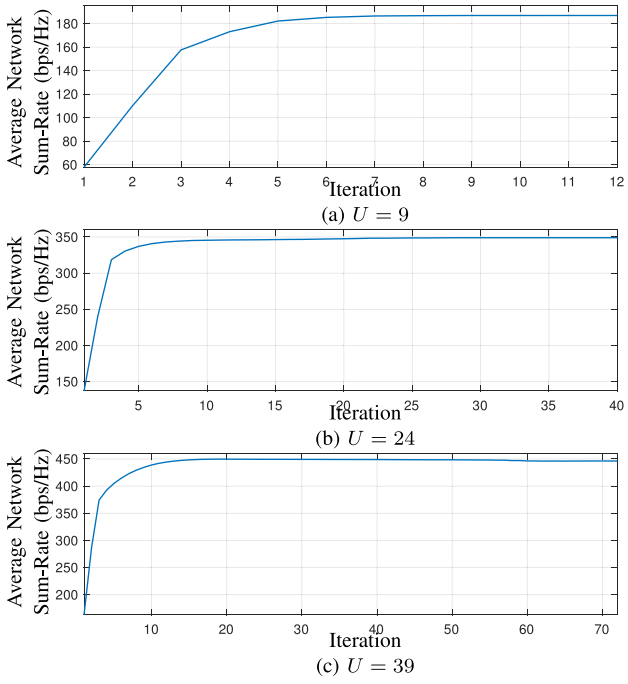


**Fig. 7.** Convergence of Average Network Sum-Rate of **JSAPA** for 9 UEs in (a), 24 UEs in (b), and 39 UEs in (c).

very limited, which increases the competition between UEs, and hence smaller number of swaps are approved.

Fig. 6 shows the average total number of iterations, at which **JSAPA** converged versus the number of UEs in the network. The total number of iterations of **JSAPA** is calculated based on the analysis presented in J as follows: total iterations = **IUSMA** iterations + **FW-DCPA** iterations + (swap operations)*(**FW-DCPA** iterations). Therefore, based on the results presented in Fig. 4 and Fig. 5, the number of iterations is expected to increase as the number of UEs in the network increases. However, this increase is more pronounced for JT-CoMP-NOMA than C-NOMA and OFDMA as the number of UEs in the network increases beyond 24. This is because the larger the UEs density, the larger the accessibility to neighboring cells' SCs, and hence the larger the matching space. For a selected number of UEs ($U = 9$, 24, and 39), the convergence behavior of the network sum-rate is presented in Fig. 7. As can be seen from Fig. 7, for a predefined error tolerance $\epsilon = 10^{-3}$, the network sum-rate increases sharply in the first few iterations, before it gradually approaches the average network sum-rate values, obtained in Fig. 3-(a).

Fig. 8(a) shows the network sum-rate for JT-CoMP-NOMA, averaged over 300 network instances, for different values of $\mu$ and $\kappa$, as per the following scenarios:

**Scenario 1:** $\mu = 3, \kappa = 2$,
**Scenario 2:** $\mu = 3, \kappa = 3$,
**Scenario 3:** $\mu = 3, \kappa = 4$,
**Scenario 4:** $\mu = 2, \kappa = 2$,

where **Scenarios 1**, **2**, **3** and **4** have a maximum clique size of $\kappa$ = 2, 3, 4 and 2, respectively. It can be observed that for **Scenarios 1**, **2** and **3**, the network sum-rate decreases as $\kappa$ increases while keeping $\mu$ constant. This can be explained as follows, as $\kappa$ increases, more UEs are allowed to be multiplexed over a single SC, thereby increasing the intra-cell interference between UEs over the same SC and consequently, the network sum-rate decreases. It is also noticed that the difference in network sum-rate is negligible in **Scenarios 2** and **3**. On the other hand, the network-sum rate in **Scenario 4** further decreases because each UE is only allowed to be assigned at most 2 SCs as opposed to the scenarios in which $\mu = 3$.

Fig. 8(b) illustrates the data rate achieved per cell-center UE in the network. Similar to the network sum-rate, the data rate of cell-center UEs are the highest in **Senario 1**, whereas it is the lowest in **Scenario 4**. Fig. 8(c), presents the data rate achieved per cell-edge UE in the network. Different observations can be made over two domains of $U$ that are $U \leq 22$ and $U > 22$. When $U \leq 22$, the number of UEs within the coverage area of each cell is less than or equal to the number of SCs $K$ in that cell. Therefore, when $\mu = 3$, most of the UEs are scheduled on 3 SCs, and thus resulting in a higher data rate than in the scenario when $\mu = 2$. It is also noticed that increasing $\kappa$ does not result in any improvement or degradation. This is because the number of UEs is still small even for $\kappa = 2$ and thus, most of the UEs can be guaranteed to be scheduled over the available spectrum, i.e., $K = 20$ SCs.

For the domain $U > 20$, as the number of UEs increases, **Scenario 1** starts to show some improvement over **Scenarios 2** and **3**. Additionally, the data rate for cell-edge UEs in **Scenario 4** continues to decrease but at a much slower rate than the in the other scenarios – allowing cell-edge UEs to achieve better data rate than in the other scenarios – as the total number of UEs in the network increases. This is because by reducing the number SCs that can be allocated to $\mu = 2$, more SCs become available for cell-edge UEs than in the case of $\mu = 3$. Furthermore, reducing $\kappa$ to 2 reduces the intra-cell interference among the UEs over a single SC, and thus cell-edge UEs are able to easily satisfy their QoS.

## 8. Conclusions

In this paper, we have studied the problem of joint sub-carrier assignment and power allocation in downlink JT-CoMP-NOMA networks, with the objective of maximizing the network sum-rate. Practical constrains, such as maximum number of UEs multiplexed over each SC, SIC decoding order, user pairing, and minimum rate requirements were considered. A generic resource allocation algorithm has been proposed for JT-CoMP-NOMA, C-NOMA and OFDMA, and its stability, convergence, complexity, and optimality properties have been analyzed. Simulation results show that by enabling JT-CoMP in NOMA, the network sum-rate and achievable data rates of cell-center UEs slightly decrease. Nevertheless, JT-CoMP-NOMA improves the achievable data rate of cell-edge UEs by 0.1 – 27.7% in comparison to C-NOMA, depending on the number of UEs in the network. Finally, the quality of the solution computed the proposed scheme was verified against a baseline solution computed by a commercial optimization package, where we were able to achieve 91.38% of the baseline solution for JT-CoMP-NOMA.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

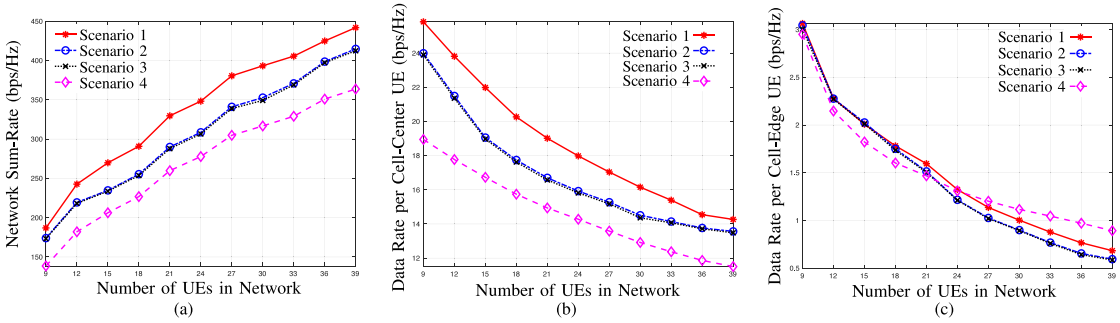


**Fig. 8.** Data rates in JT-CoMP-NOMA network for different values of $\mu$ and $\kappa$ vs. the total number of UEs in the network.

## Acknowledgments

This work was supported and funded by Kuwait University Research Grant No. EO-05/19, EO-08/18 and the Kuwait Foundation for the Advancement of Sciences (KFAS), under Project Code PN17-15EE-02 and P314-35EO-01.

## Appendix A

In this Appendix we elaborate on the concept of "hole". Assume the players initiating a *swap operation* are $u$, $H_{u'}$, $k$ and $k'$, such that $k \in \Psi(u)$ and $H_{u'} \in \Psi(k')$, i.e., $|\Psi(k')| < \kappa$.[13] Therefore, after the swap operation, a swap matching $\Psi^{uk}_{H_{u'}k'} = \Psi \backslash \{(u,k)\} \cup \{(u,k')\}$ results such that $k' \in \Psi(u)$, $H_{u'} \in \Psi(k)$, i.e., $|\Psi(k)| < \kappa$, and $k \notin \Psi(u)$, $|\Psi(k')| \leq \kappa$.[14] On the other hand, consider a scenario where the players initiating a swap operation are $u$, $H_{u'}$, $H_k$, and $k'$, such that $H_k \in \Psi(u)$, i.e., $|\Psi(u)| < \mu$, and $H_{u'} \in k'$, i.e., $|\Psi(k')| < \kappa$. In this scenario, there are two holes, namely $H_{u'}$ and $H_k$. The latter indicates that SC $k'$ has an empty spot for a UE to be multiplexed, and the former indicates that UE $u$ can be allocated one more SC. Subsequently, a swap matching $\Psi^{uH_k}_{H_{u'}k'} = \Psi \cup \{(u,k')\}$ is defined such that $k' \in \Psi(u)$, which occurs after the swap operation.

## Appendix B

**Lemma 1.** *The computational complexity of **RICSF** is upper bounded by* $O\left(\kappa|\mathbb{V}_k^b|^\kappa\right)$.

**Proof.** Let the time complexity of **RICSF** be denoted by $T(q)$, where $q$ is the size of the clique $Q$ to be increased. Notice that the recursion at line 7 in **Algorithm 2** calls **RICSF** with $(q+1)$, and thus let its time complexity be $T(q+1)$. Further, the number of vertices $\hat{v}$ (line 5) that could be added to $Q$ to form a new clique $Q^*$ of size $(q+1)$ is equal to total number of vertices in the graph, i.e., $|\mathbb{V}_k^b|$, in the worst case scenario. Consequently, in the worst case scenario, line 7 in **Algorithm 2** is called $|\mathbb{V}_k^b|$ times and thus, the time complexity of **RICSF** can be written as $T(q) = |\mathbb{V}_k^b|T(q+1)$. Following the same logic, the time complexity of **RICSF** called with $(q+1)$ is written as $T(q+1) = |\mathbb{V}_k^b|T(q+2)$; with $(q+2)$, $T(q+2) = |\mathbb{V}_k^b|T(q+3)$; with $(q+3)$, $T(q+3) = |\mathbb{V}_k^b|T(q+4)$; and so on. Applying the substitution method,

$$T(q) = |\mathbb{V}_k^b|T(q+1) = |\mathbb{V}_k^b|^2 T(q+2) = |\mathbb{V}_k^b|^3 T(q+3) = \cdots = |\mathbb{V}_k^b|^c T(q+c),$$

where $c$ is a constant. The substitution stops at some point where no more recursion occurs and the algorithm terminates. This is achieved by what is called the base condition when $q = \kappa$ (line 6 in **Algorithm 2**). Accordingly, the recursion stops when $T(q+c) = T(\kappa)$ and this is obtained for $c = \kappa - q$. Substituting the value of $c$ we get $T(q) = |\mathbb{V}_k^b|^{\kappa-q} T(\kappa) = \kappa|\mathbb{V}_k^b|^{\kappa-q}$. As a result, the computational complexity of **RICSF** is upper bounded by $O(\kappa|\mathbb{V}_k^b|^{\kappa-q})$. □

## Appendix C

**Lemma 2.** *The computational complexity of **K-UGFA** is upper bounded by* $O(2BK|\mathbb{V}_k^b| + BK|\mathbb{L}_k^b| + \kappa BK|\mathbb{V}_k^b|^{\kappa-q+1})$.

**Proof.** The complexity of **K-UGFA** depends on the clique formation process (Step 2) in **Algorithm 1**, which consists of constructing the graphs and finding the cliques using **RICSF**. In addition to the complexity of **RICSF**, which according to Lemma 1 is $O\left(\kappa|\mathbb{V}_k^b|^{\kappa-q}\right)$, we need to also consider the complexity of constructing the graphs (line 4) and forming 1-cliques (line 6). The complexity for constructing graph $\mathbb{G}_k^b$ is $O(|\mathbb{V}_k^b| + |\mathbb{L}_k^b|)$. On the other hand, it takes at most $|\mathbb{V}_k^b|$ iterations to form 1-cliques. In addition, **RICSF** is called $|\mathbb{V}_k^b|$ times in **Algorithm 1** (line 7). Eventually, note that cliques are formed for all SCs in all BSs in the network. Thus, the computational complexity of **K-UGFA** is upper bounded by $BK \cdot O(2|\mathbb{V}_k^b| + |\mathbb{L}_k^b| + \kappa|\mathbb{V}_k^b \parallel \mathbb{V}_k^b|^{\kappa-q}) \simeq O(2BK|\mathbb{V}_k^b| + BK|\mathbb{L}_k^b| + \kappa BK|\mathbb{V}_k^b|^{\kappa-q+1})$. □

## Appendix D

**Lemma 3.** *The **SMA** algorithm results in a matching $\Psi^*$ that is a 2ES matching.*

**Proof.** Recall that due to the existence of externalities, the traditional notion of stability no longer exists. On the contrary, the proposed **SMA** results in a two-sided exchange stable matching where no swap-blocking pairs exist and all players are satisfied with their matchings. Assume that the resulting matching $\Psi^*$ contains a swap-blocking pair $(u,u')$ where $k \in \Psi^*(u)$ and $k' \in \Psi^*(u')$, such that $\forall i \in \{u,u',k,k'\}, U_i\left(\Psi^{uk}_{u'k'}\right) \geq U_i(\Psi^*)$ and $\exists i \in \{u,u',k,k'\}, U_i\left(\Psi^{uk}_{u'k'}\right) > U_i(\Psi^*)$. However, according to **SMA** in **Algorithm 4**, it stops when no swap-blocking pairs are found; in other words, all swap-blocking pairs are eliminated. This contradicts with the initial assumption that $\Psi^*$ contains a swap-blocking pair $(u,u')$. Moreover, according to Definition 5, a matching that does not contain any swap-blocking pairs is a 2ES matching. Hence, the resulting matching $\Psi^*$ of **SMA** is *2ES*. □

## Appendix E

**Lemma 4.** *The proposed **IUSMA** converges to an initial matching $\Psi$ in a finite number of iterations.*

[13] Note that $H_u \in \Psi(k')$ does not mean that $k'$ is matched to a hole. Instead, it implies that $k'$ is not fully matched, i.e., $|\Psi(k')| < \kappa$.

[14] Note that before the swap, the number matchings of $k'$ was strictly less than $\kappa$ because of the hole. However, after the swap, the number of matchings of $k'$ could be either equal to $\kappa$ (if $k'$ had only one H), or could still be less than $\kappa$ (if $k'$ had more than one H).

**Proof.** In each iteration, an unmatched UE $u \in \mathcal{Z}$ proposes to its most preferred SC that has never rejected it before. Therefore, as the number of iterations increases, the set of choices available to UE $u$ in its preference list $\mathcal{P}_u$ becomes smaller. There are $K$ SCs in each BS; therefore, the size of the preference list of each cell-center UE is at most $K$. For cell-edge UEs, the number of available SCs depends on the number of BSs in their range and the number of coordinated BSs $\tau$. Nevertheless, according to Remark 4, the preference lists are incomplete and therefore their size is limited. Hence, the total number of iterations is also limited. This confirms that **IUSMA** converges to an initial matching $\Psi$ with a finite number of iterations. □

## Appendix F

**Lemma 5.** *The proposed **SMA** algorithm converges to a 2ES matching $\Psi^*$ in a finite number of iterations.*

**Proof.** The convergence of **SMA** mainly depends on the swap matching process (Step 2 of **Algorithm 4**). According to Definition 4, the utilities of all players $\forall i \in \{u, u', k, k'\}$ in matching $\Psi$ that are participating in a swap satisfy: $U_i\left(\Psi^{uk}_{u'k'}\right) \geq U_i(\Psi)$, in which at least one of the players' utility in $\Psi^{uk}_{u'k'}$ is strictly greater than that of its utility in $\Psi$, i.e., $U_i\left(\Psi^{uk}_{u'k'}\right) > U_i(\Psi)$. Therefore, the total network sum-rate increases after each swap operation: $U_{total}\left(\Psi^{uk}_{u'k'}\right) - U_{total}(\Psi) > 0$. Note that due to the finite number of players, i.e., UEs and SCs, the number of potential swap-blocking pairs is finite and hence, the number of swap operations is also finite. Moreover, the network sum-rate has an upper bound due to limited spectrum resources, and thus there exists a swap operation after which no more swap-blocking pairs are found that can further increase the network sum-rate. Accordingly, **SMA** converges to a 2ES matching $\Psi^*$ (Lemma 3) in a finite number of iterations. □

## Appendix G

**Lemma 6.** *The complexity of the **IUSMA** is $O(TK^2 + EC^2)$, where $T = |\mathcal{T}|, K = |\mathcal{K}|, E = |\mathcal{E}|$, and $C$ are the number of cell-center UEs, the number of cell-edge UEs, the number of SCs and the number of all possible subsets of SCs that cell-edge UEs can be matched to, respectively.*

**Proof.** The computational complexity of **IUSMA** depends on the construction of preference lists (Step 1) and the matching process (Step 2) in **Algorithm 3**. In the preference lists construction phase, each cell-center UE obtains its preference list over $K$ SCs – from the single BS that it is associated with – in which the complexity is $O(TK^2)$. On the other hand, each cell-edge UE obtains its preference list over $C$ possible subsets of SCs – from all BSs within its range – in which the complexity is $O(EC^2)$. Therefore, the overall complexity of preference lists construction is $O(TK^2 + EC^2)$. Furthermore, in the matching process, each cell-center UE proposes at most $K$ times, whereas each cell-edge UE proposes at most $C$ times, resulting in an overall complexity of $O(TK + EC)$ for the matching process. Consequently, this results in a total computational complexity of $O(TK^2 + EC^2)$ [60] for **IUSMA**. □

## Appendix H

**Lemma 7.** *All local maxima of network sum-rate correspond to a 2ES matching.*

**Proof.** Consider a matching $\Psi$ with a network sum-rate that is a local maximum value. If $\Psi$ is not a 2ES matching, then there exist a swap-blocking pair such that when a swap operation is approved the network sum-rate increases, as per Lemma 5. However, this contradicts the assumption that $\Psi$ is a local maximum, and thence $\Psi$ is a 2ES matching [18,20,47]. □

Nevertheless, it is important to note that not all 2ES matchings resulting from **SMA** are local maxima of the network sum-rate. To see this, given the players $\{u, u', k, k'\}$ participating in a swap, $u'$ may not approve the swap with $u$ because it decreases its utility despite the fact that it increases all other players' utilities. Approving such a swap would indeed increase the network sum-rate but at the expense of a weaker stability [18,20,47].

## Appendix I

**Proposition 1.** *The computational complexity of **FW-DCPA** is upper bounded by $O(IBKU^3)$.*

**Proof.** The computational complexity of solving (31) for a network where $K = 1$ is $O(U^3)$ [26,52]. Since in our network each BS has $K > 1$ SCs, the complexity becomes $K \cdot O(U^3) \simeq O(KU^3)$. Eventually, as there are $B$ BSs in the system, where each BS has $K$ SCs, the complexity of solving (31) is $B \cdot O(KU^3) \simeq O(BKU^3)$. Accordingly, given the total number of iterations $I$ of **FW-DCPA**, its computational complexity is given by $O(IBKU^3)$. □

## Appendix J

**Proposition 2.** *The computational complexity of the proposed **JSAPA** scheme is upper bounded by $O(2BK|\mathbb{V}^b_k| + BK|\mathbb{K}^b_k| + \kappa BK|\mathbb{V}^b_k|^{\kappa-q+1} + TK^2 + EC^2 + IBKU^3 + SIBKU^3)$.*

**Proof.** The computational complexity of **JSAPA** depends on the complexity of the algorithms it comprises. Therefore, according to **Algorithm 6**, the complexity of **JSAPA** is obtained as follows, **JSAPA** = **K-UGFA** + **IUSMA** + **FW-DCPA** + (number of swaps in **SMA**)***FW-DCPA**. The complexity of **K-UGFA** is $O(2BK|\mathbb{V}^b_k| + BK|\mathbb{L}^b_k| + \kappa BK|\mathbb{V}^b_k|^{\kappa-q+1})$, **IUSMA** is $O(TK^2 + EC^2)$, and **FW-DCPA** is $O(IBKU^3)$ according to Lemma 2, Lemma 6 and Proposition 1, respectively. In addition, let $S$ be the total number of swap operations performed in **SMA**. Consequently, the computational complexity of **JSAPA** is upper bounded by $O(2BK|\mathbb{V}^b_k| + BK|\mathbb{K}^b_k| + \kappa BK|\mathbb{V}^b_k|^{\kappa-q+1} + TK^2 + EC^2 + IBKU^3 + SIBKU^3)$. □

## Appendix K

**Proposition 3.** *The proposed **JSAPA** is guaranteed to converge in a finite number of iterations.*

**Proof.** The convergence of **JSAPA** depends on the convergence of its sub-algorithms **K-UGFA**, **IUSMA**, **SMA** and **FW-DCPA**. Moreover, the convergence of **K-UGFA** depends on the clique formation process (Step 2) in **Algorithm 1**, which consists of constructing the graphs and finding the cliques using the recursive function **RICSF**. **Algorithm 2** takes as an input a clique $Q$ of size $q$. Each call of the **RICSF** function returns a clique $Q^*$ of size $q + 1$. The recursive call of **RICSF** stops at some point where no more recursions occur, and the algorithm terminates. This is achieved by what is called the base condition when $q = \kappa$ (line 6 in **Algorithm 2**). Subsequently, **RICSF** is guaranteed to converge. As discussed in Lemma 4 (Appendix E) and Lemma 5 (Appendix F), the proposed **IUSMA** and **SMA** algorithms converge in a finite number of iterations.

Finally, based on the discussion in Section 5, in **Algorithm 5**, the solution $\mathbf{p}^{(t+1)}$ is generated as the optimal solution of problem (31) using the solution of the previous iteration, i.e., p(t). Furthermore, according to [24,49], the solution $\mathbf{p}^{(t+1)}$ at the $(t+1)$ iteration is always better than $\mathbf{p}^{(t)}$ at the $t$th iteration. This implies that the objective function $V(\mathbf{p})$ improves in each iteration. Additionally, the set of constraints in (31) is compact and thus, by Cauchy theorem, the sequence of generated $\mathbf{p}^{(t)}$ always converges [49]. Hence, JSAPA is guaranteed to converge in a finite number of iterations. □

## References

[1] Z. Ding, X. Lei, G.K. Karagiannidis, R. Schober, J. Yuan, V.K. Bhargava, A survey on non-orthogonal multiple access for 5g networks: Research challenges and future trends, IEEE J. Sel. Areas Commun. 35 (10) (2017) 2181–2195.
[2] Y. Feng, S. Yan, Z. Yang, N. Yang, J. Yuan, Beamforming design and power allocation for secure transmission with NOMA, IEEE Trans. Wireless Commun. 18 (5) (2019) 2639–2651.
[3] X. Sun, S. Yan, N. Yang, Z. Ding, C. Shen, Z. Zhong, Short-packet downlink transmission with non-orthogonal multiple access, IEEE Trans. Wireless Commun. 17 (7) (2018) 4550–4564.
[4] Y. Yuan, Z. Yuan, L. Tian, 5G non-orthogonal multiple access study in 3GPP, IEEE Commun. Mag. 58 (7) (2020) 90–96.
[5] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, K. Higuchi, Non-orthogonal multiple access (NOMA) for cellular future radio access, in: Proc. of IEEE 77th Vehicular Technology Conference, VTC Spring, 2013, pp. 1–5.
[6] J. Zeng, T. Lv, R.P. Liu, X. Su, M. Peng, C. Wang, J. Mei, Investigation on evolving single-carrier NOMA into multi-carrier NOMA in 5G, IEEE Access 6 (2018) 48268–48288.
[7] W. Shin, M. Vaezi, B. Lee, D.J. Love, J. Lee, H.V. Poor, Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges, IEEE Commun. Mag. 55 (10) (2017) 176–183.
[8] Z. Yang, C. Pan, W. Xu, Y. Pan, M. Chen, M. Elkashlan, Power control for multi-cell networks with non-orthogonal multiple access, IEEE Trans. Wireless Commun. 17 (2) (2018) 927–942.
[9] Y. Fu, Y. Chen, C.W. Sung, Distributed power control for the downlink of multi-cell NOMA systems, IEEE Trans. Wireless Commun. 16 (9) (2017) 6207–6220.
[10] F. Qamar, K.B. Dimyati, M.N. Hindia, K.A.B. Noordin, A.M. Al-Samman, A comprehensive review on coordinated multi-point operation for LTE-A, Comput. Netw. 123 (2017) 19–37.
[11] M. Ali, E. Hossain, D.I. Kim, Coordinated multipoint transmission in downlink multi-cell NOMA systems: Models and spectral efficiency performance, IEEE Wirel. Commun. 25 (2) (2018) 24–31.
[12] Q. Guo, C.W. Sung, Y. Chen, C.S. Chen, Power control for coordinated NOMA downlink with cell-edge users, in: Proc. of IEEE Wireless Communications and Networking Conference, WCNC, 2018, pp. 1–6.
[13] M. Ali, E. Hossain, A. Al-Dweik, D.I. Kim, Downlink power allocation for CoMP-NOMA in multi-cell networks, IEEE Trans. Commun. 66 (9) (2018) 3982–3998.
[14] X. Sun, N. Yang, S. Yan, Z. Ding, D.W.K. Ng, C. Shen, Z. Zhong, Joint beamforming and power allocation in downlink NOMA multiuser MIMO networks, IEEE Trans. Wireless Commun. 17 (8) (2018) 5367–5381.
[15] Y. Al-Eryani, E. Hossain, D.I. Kim, Generalized coordinated multipoint (GCoMP)-enabled NOMA: Outage, capacity, and power allocation, IEEE Trans. Commun. 67 (11) (2019) 7923–7936.
[16] Y. Dai, L. Lyu, NOMA-enabled CoMP clustering and power control for green Internet of things networks, IEEE Access 8 (2020) 90109–90117.
[17] S. Bayat, Y. Li, L. Song, Z. Han, Matching theory: Applications in wireless communications, IEEE Signal Process. Mag. 33 (6) (2016) 103–122.
[18] J. Zhao, Y. Liu, K.K. Chai, A. Nallanathan, Y. Chen, Z. Han, Spectrum allocation and power control for non-orthogonal multiple access in hetnets, IEEE Trans. Wireless Commun. 16 (9) (2017) 5825–5837.
[19] S.A. Kazmi, N.H. Tran, T.M. Ho, A. Manzoor, D. Niyato, C.S. Hong, Coordinated device-to-device communication with non-orthogonal multiple access in future wireless cellular networks, IEEE Access 6 (2018) 39860–39875.
[20] B. Di, L. Song, Y. Li, Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks, IEEE Trans. Wireless Commun. 15 (11) (2016) 7686–7698.
[21] M. Baidas, M. Al-Mubarak, E. Alsusa, M.K. Awad, Joint subcarrier assignment and global energy-efficient power allocation for energy-harvesting two-tier downlink NOMA hetnets, IEEE Access 7 (2019) 163556–163577.
[22] M.S. Ali, H. Tabassum, E. Hossain, Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems, IEEE Access 4 (2016) 6325–6343.
[23] Z. Ding, P. Fan, H.V. Poor, Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions, IEEE Trans. Veh. Technol. 65 (8) (2016) 6010–6023.
[24] M.B. Shahab, M. Irfan, M.F. Kader, S. Young Shin, User pairing schemes for capacity maximization in non-orthogonal multiple access systems, Wireless Commun. Mob. Comput. 16 (17) (2016) 2884–2894.
[25] R. Singh, Sub-channel assignment and resource scheduling for non-orthogonal multiple access (NOMA) in downlink coordinated multi-point systems, in: Proc. of IEEE 20th Conference on Innovations in Clouds, Internet and Networks, ICIN, 2017, pp. 17–22.
[26] Z. Liu, G. Kang, L. Lei, N. Zhang, S. Zhang, Power allocation for energy efficiency maximization in downlink CoMP systems with NOMA, in: Proc. of IEEE Wireless Communications and Networking Conference, WCNC, 2017, pp. 1–6.
[27] Y. Tian, A.R. Nix, M. Beach, On the performance of opportunistic NOMA in downlink CoMP networks, IEEE Commun. Lett. 20 (5) (2016) 998–1001.
[28] L. Salaün, C.S. Chen, M. Coupechoux, Optimal Joint Subcarrier and Power Allocation in NOMA is Strongly NP-Hard, in: Proc. of IEEE International Conference on Communications, ICC, 2018, pp. 1–7.
[29] C. Bron, J. Kerbosch, Algorithm 457: finding all cliques of an undirected graph, Commun. ACM 16 (9) (1973) 575–577.
[30] H. Tuy, Convex Analysis and Global Optimization, Kluwer Academic, 1998.
[31] F. Marguerite, W. Philip, An algorithm for quadratic programming, Nav. Res. Logist. Q. 3 (1–2) (1956) 95–110.
[32] M. K. Awad, M. W. Baidas, A. A. El-Amine, Optimal resource allocation for joint transmission CoMP-eenabled NOMA networks: A benchmark implementation, in: Proc. of IEEE 36th National Radio Science Conference, NRSC, 2019, pp. 249–258.
[33] J. Lee, Y. Kim, H. Lee, B.L. Ng, D. Mazzarese, J. Liu, W. Xiao, Y. Zhou, Coordinated multipoint transmission and reception in LTE-advanced systems, IEEE Commun. Mag. 50 (11) (2012) 44–50.
[34] Z. Li, Y. Liu, Y. Zhang, P. Wang, C. Tian, X. Sha, Hybrid coordinated strategy of downlink coordinated multi-point transmission, J. Syst. Eng. Electron. 26 (5) (2015) 916–923.
[35] J. Pastor-Perez, F. Riera-Palou, G. Femenias, FFR-aided coordinated multipoint transmission in downlink multicell MIMO-OFDMA networks, in: 2015 IEEE 82nd Vehicular Technology Conference, VTC2015-Fall, 2015, pp. 1–7.
[36] Q. Zhang, C. Yang, Transmission mode selection for downlink coordinated multipoint systems, IEEE Trans. Veh. Technol. 62 (1) (2013) 465–471.
[37] S. Sen, N. Santhapuri, R.R. Choudhury, S. Nelakuditi, Successive interference cancellation: A back-of-the-envelope perspective, in: Proc. of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks, in: Hotnets-IX, ACM, New York, NY, USA, 2010, pp. 17:1–17:6, (Online), Available: http://doi.acm.org/10.1145/1868447.1868464.
[38] S.R. Islam, N. Avazov, O.A. Dobre, K.-S. Kwak, Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges, IEEE Commun. Surv. Tutor. 19 (2) (2017) 721–742.
[39] J. Zhu, J. Wang, Y. Huang, S. He, X. You, L. Yang, On optimal power allocation for downlink non-orthogonal multiple access systems, IEEE J. Sel. Areas Commun. 35 (12) (2017) 2744–2757.
[40] D.W.K. Ng, E.S. Lo, R. Schober, Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas, IEEE Trans. Wireless Commun. 11 (9) (2012) 3292–3304.
[41] W. Cai, C. Chen, L. Bai, Y. Jin, J. Choi, Subcarrier and power allocation scheme for downlink OFDM–NOMA systems, IET Signal Process. 11 (1) (2017) 51–58.
[42] L. Dai, B. Wang, Y. Yuan, S. Han, C. l. I., Z. Wang, Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends, IEEE Commun. Mag. 53 (9) (2015) 74–81.
[43] J. McGee, T. Sammut-Bonnici, Network externalities, in: Wiley Encyclopedia of Management, Wiley Online Library, 2015, pp. 1–5.
[44] V. Bansal, A. Agrawal, V.S. Malhotra, Polynomial time algorithm for an optimal stable assignment with multiple partners, Theoret. Comput. Sci. 379 (3) (2007) 317–328.
[45] K. Iwama, S. Miyazaki, Y. Morita, D. Manlove, Stable marriage with incomplete lists and ties, in: J. Wiedermann, P. van Emde Boas, M. Nielsen (Eds.), Autom. Lang. Programming (1999) 443–452.
[46] F. Echenique, J. Oviedo, A Theory of Stability in Many-to-Many Matching Markets, (1185) Caltech SS Working Paper, 2004.
[47] E. Bodine-Baron, C. Lee, A. Chong, B. Hassibi, A. Wierman, Peer effects and stability in matching markets, in: G. Persiano (Ed.), Algorithmic Game Theory (2011) 117–129.
[48] M. Baccara, A. İmrohoroğlu, A.J. Wilson, L. Yariv, A field study on matching with network externalities, Amer. Econ. Rev. 102 (5) (2012) 1773–1804.
[49] D. Gale, L.S. Shapley, College admissions and the stability of marriage, Amer. Math. Monthly 69 (1) (1962) 9–15.
[50] J.M. Kumpula, M. Kivelä, K. Kaski, J. Saramäki, Sequential algorithm for fast clique percolation, Phys. Rev. E 78 (2) (2008) 026109.
[51] B. Dutta, J. Massó, Stability of matchings when individuals have preferences over colleagues, J. Econom. Theory 75 (2) (1997) 464–475.
[52] H.H. Kha, H.D. Tuan, H.H. Nguyen, Fast global optimal power allocation in wireless networks by local DC programming, IEEE Trans. Wireless Commun. 11 (2) (2012) 510–515.
[53] F. Fang, H. Zhang, J. Cheng, V.C. Leung, Energy efficiency of resource scheduling for non-orthogonal multiple access (NOMA) wireless network, in: Proc. of IEEE International Conference on Communications, ICC, 2016, pp. 1–5.
[54] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge university press, 2004.
[55] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, 2014, http://cvxr.com/cvx.
[56] A. Goldsmith, Wireless Communications, Cambridge University Press, 2005.
[57] D.R. Smith, Digital Transmission Systems, third ed., Springer Science & Business Media, LLC, 2004.

[58] M. Tawarmalani, N.V. Sahinidis, A polyhedral branch-and-cut approach to global optimization, Math. Programm. 103 (2) (2005) 225–249.
[59] G.D. Corporation, General algebraic modeling system (GAMS), 2020, (Online), Available:https://www.gams.com.
[60] B. Di, S. Bayat, L. Song, Y. Li, Radio resource allocation for downlink non-orthogonal multiple access (NOMA) networks using matching theory, in: Proc. of IEEE Global Communications Conference, GLOBECOM, 2015, pp. 1–6.

**Mohamad Khattar Awad** (S'02, M'09, SM'17), earned the B.A.Sc. in electrical and computer engineering (communications option) from the University of Windsor, Ontario, Canada, in 2004 and the M.A.Sc. and Ph.D. in electrical and computer engineering from the University of Waterloo, Ontario, Canada, in 2006 and 2009, respectively.

From 2004 to 2009 he was a research assistant in the Broadband Communications Research Group (BBCR), University of Waterloo. In 2009 to 2012, he was an Assistant Professor of Electrical and Computer Engineering at the American University of Kuwait. Since 2012, he has been with Kuwait University, where he currently is an Associate Professor of Computer Engineering.

Dr. Awad's research interest includes wireless and wired communications, software-defined networks resource allocation, wireless networks resource allocation, and acoustic vector-sensor signal processing. He received the Ontario Research & Development Challenge Fund Bell Scholarship in 2008 and 2009, the University of Waterloo Graduate Scholarship in 2009, and a fellowship award from the Dartmouth College, Hanover, NH in 2011. In 2015 and 2017, he received the Kuwait University Teaching Excellence Award and Best Young Researcher Award, respectively. Dr. Awad serves on the editorial board of the IEEE Transactions on Green Communications and Networking (TGCN).

**Mohammed W. Baidas** (M'05-SM'17) received the B.Eng. (Hons.) degree in communication systems engineering from the University of Manchester, Manchester, U.K., in 2005, the M.Sc. degree (with distinction) in wireless communications engineering from the University of Leeds, Leeds, U.K., in 2006, the M.S. degree in electrical engineering from the University of Maryland, College Park, MD, USA, in 2009, and the Ph.D. degree in electrical engineering from Virginia Tech, Blacksburg, VA, USA, in 2012. He was a Visiting Researcher with the University of Manchester in the Academic Years of 2015/2016 and 2018/2019. He is currently an Associate Professor with the Department of Electrical Engineering, Kuwait University, Kuwait, where he has been on the faculty since May 2012. He is also a frequent reviewer for several IEEE journals and international journals and conferences, with over 70 publications. His research interests include resource allocation and management in cognitive radio systems, game theory, cooperative communications and networking, and green and energy-harvesting networks. He also serves as a technical program committee member for various IEEE and international conferences. He was a recipient of the Outstanding Teaching Award of Kuwait University for the academic year of 2017/2018.

**Ahmad El-Amine** (M'18) received the B.Eng. (high distinction) degree in computer and communication engineering from Notre Dame University, Louaize, Lebanon, in 2016. He is currently a research assistant at the Department of Computer Engineering, Kuwait University, Kuwait. His research interests include resource allocation and management in wireless cellular networks, game theory, and optimization.