Power and Channel Allocation for Non-Orthogonal Multiple Access in 5G Systems: Tractability and Computation

Lei Lei, Di Yuan, Senior Member, IEEE, Chin Keong Ho, Member, IEEE, and Sumei Sun, Fellow, IEEE

Abstract—A promising multi-user access scheme, non-orthogonal multiple access (NOMA) with successive interference cancellation (SIC), is currently under consideration for 5G systems. NOMA allows more than one user to simultaneously access the same frequency-time resource and separates multi-user signals by SIC. These render resource optimization in NOMA different from orthogonal multiple access. We provide theoretical insights and algorithmic solutions to jointly optimize power and channel allocation in NOMA. We mathematically formulate NOMA resource allocation problems, and characterize and analyze the problems’ tractability under a range of constraints and utility functions. For tractable cases, we provide polynomial-time solutions for global optimality. For intractable cases, we prove the NP-hardness and propose an algorithmic framework combining Lagrangian duality and dynamic programming to deliver near-optimal solutions. To gauge the performance of the solutions, we also provide optimality bounds on the global optimum. Numerical results demonstrate that the proposed algorithmic solution can significantly improve the system performance in both throughput and fairness over orthogonal multiple access as well as over a previous NOMA resource allocation scheme.

Index Terms—Non-orthogonal multiple access, resource allocation, successive interference cancellation, 5G.

I. INTRODUCTION

ORTHOGONAL multi-user access (OMA) techniques are used in 4G long term evolution (LTE) and LTE-Advanced (LTE-A) networks, e.g., orthogonal frequency division multiple access (OFDMA) for downlink and single-carrier frequency division multiple access (SC-FDMA) for uplink [1], [2]. In OMA, within a cell, each user has exclusive access to the allocated resource blocks. Thus, each subchannel or subcarrier can only be utilized by at most one user in every time slot. OMA avoids intra-cell interference, and enables single-user detection/decoding and simple receiver design. However, by its nature, orthogonal channel access is becoming a limiting factor of spectrum efficiency.

In the coming decade, mobile data traffic is expected to grow thousand-fold [3], [4]. Accordingly, the network capacity must dramatically increase for 5G systems. Capacity scaling for 5G is enabled by a range of techniques and schemes, e.g., cell densification, utilization of unlicensed spectrum, and advanced radio access schemes. New multi-user access schemes have been investigated as potential alternatives to OFDMA and SC-FDMA [3]–[5]. A promising scheme is the so-called non-orthogonal multiple access (NOMA) with successive interference cancellation (SIC) [6]. Unlike interference-avoidance multiple access schemes, e.g., OFDMA, multiple users in NOMA can be assigned to the same frequency-time resource so as to improve spectrum efficiency [6]. On one hand, this results in intra-cell interference among the multiplexed users. On the other hand, some of the interfering signals in NOMA can be eliminated by multi-user detection (MUD) with SIC at the receiver side. To enable this process, more advanced receiver design and interference management techniques are considered for 5G networks [4].

A. Related Works

From an information theory perspective, under the assumption of simultaneous multi-user transmission via superposition coding with SIC, capacity region and duality analysis have been studied in [7]. The authors of [8] have provided an analysis of implementing interference cancellation in cellular systems. Towards future 5G communication systems, some candidate access schemes are under investigation in recent research activities, e.g., sparse code multiple access (SCMA) [3], and NOMA [6].

In [9], the authors studied the capacity region for NOMA. In [10], by assuming predefined user groups for each subchannel, a heuristic algorithm for NOMA power allocation in downlink has been proposed, and system-level simulations have been conducted. In [11], the authors address various implementation issues of NOMA. The authors in [12] considered a sum-rate utility maximization problem for dynamic NOMA resource allocation. In [13], outage performance of NOMA has been evaluated. The authors derived the ergodic sum rate and outage probability to demonstrate the superior performance of NOMA with fixed power allocation. In [14],
fairness considerations and a max-min fairness problem for NOMA have been addressed. The fairness in NOMA can be improved via using and adapting the so-called power allocation coefficients. For uplink NOMA, the authors in [15] provided a suboptimal algorithm to solve an uplink scheduling problem with fixed transmission power. In [16], a weighted multi-user scheduling scheme is proposed to balance the total throughput and the cell-edge user throughput. In [17], the authors proposed a greedy-based algorithm to improve the throughput in uplink NOMA. In [18], the authors studied and evaluated user grouping/pairing strategies in NOMA. It has been shown that, from the outage probability perspective, it is preferable to multiplex users of large gain difference on the same subcarrier. In [19], the authors address subcarrier allocation and power assignment in downlink NOMA, with the objective of balancing the throughput with the number of scheduled users. The solution approach uses matching theory, by first assuming equal power split and applying a user-subchannel matching algorithm that converges to a stable matching, followed by a water-filling phase for power allocation. In [20], a monotonic optimization method is developed for NOMA subcarrier and power allocation. The method potentially approaches the global optimum, at the cost of high complexity in the number of users per subcarrier. We remark that there are other setups of SIC than that considered in NOMA. An example is the interference channel in which common information is transmitted for partial interference cancellation, for which Etkin et al. [21] provided an analysis of the resulting capacity region and trade-off from an information theory perspective.

We note that in some studies of NOMA (e.g., [10], [14], [18]), the aspect of subcarrier or subchannel allocation was simplified or not addressed. However, to fully reach the potential of NOMA, user-subchannel allocation is of high significance due to fading. Assuming uniform subchannels or using fixed rules for subchannel allocation may result in significant performance loss. The importance of subchannel allocation is evidenced by the growing interest in explicitly taking this aspect into account in some recent works [12], [19], [20].

Apart from investigation of NOMA performance in cellular networks, from an optimization perspective, the complexity and tractability analysis of NOMA resource allocation is of significance. Here, tractability for an optimization problem refers to whether or not any polynomial-time algorithm can be expected to find the global optimum [1]. Tractability results for resource allocation in OMA and interference channels have been investigated in a few existing works, e.g., [1], [22], [23] for OFDMA, [2] for SC-FDMA, and [23], [24] for interference channel. For NOMA, to the best of our knowledge, no such study is available in the existing literature.

**A. Basic Notation**

We consider a downlink cellular system with a base station (BS) serving $K$ users. The overall bandwidth $B$ is divided into $N$ subchannels, each with bandwidth $B/N$. We note that in some studies of NOMA (e.g., [10], [14], [18]), the aspect of subcarrier or subchannel allocation was simplified or not addressed. However, to fully reach the potential of NOMA, user-subchannel allocation is of high significance due to fading. Assuming uniform subchannels or using fixed rules for subchannel allocation may result in significant performance loss. The importance of subchannel allocation is evidenced by the growing interest in explicitly taking this aspect into account in some recent works [12], [19], [20].

Apart from investigation of NOMA performance in cellular networks, from an optimization perspective, the complexity and tractability analysis of NOMA resource allocation is of significance. Here, tractability for an optimization problem refers to whether or not any polynomial-time algorithm can be expected to find the global optimum [1]. Tractability results for resource allocation in OMA and interference channels have been investigated in a few existing works, e.g., [1], [22], [23] for OFDMA, [2] for SC-FDMA, and [23], [24] for interference channel. For NOMA, to the best of our knowledge, no such study is available in the existing literature.

**B. Contributions**

In spite of the existing literature of performance evaluation for NOMA, there is lack of a systematic approach for NOMA resource allocation from a mathematical optimization point of view. The existing resource allocation approaches for NOMA are typically carried out with fixed power allocation [13], [15], [17], predefined user set for subchannels [10], or parameter tuning to improve performance, e.g., updating power allocation coefficients [14]. Moreover, compared with OMA, NOMA allows multi-user sharing on the same subchannel, thus provides an extra dimension to influence the performance in throughput and fairness. However, how to balance these two key performance aspects in power and channel allocation is largely yet not studied in the literature. In addition, little is known on the computational complexity and tractability of NOMA resource allocation.

In this paper, the solutions of joint channel and power allocation for NOMA are subject to systematic optimization, rather than using heuristics or ad-hoc methods. To this end, we formulate, analyze, and solve the power and channel optimization problem for downlink NOMA systems, taking into account practical considerations of fairness and SIC. We present the following contributions. First, for maximum weighted-sum-rate (WSR) and sum-rate (SR) utilities, we formulate the joint power and channel allocation problems (JPCAP) mathematically. Second, we prove the NP-hardness of JPCAP with WSR and SR utilities. Third, we identify tractable cases for JPCAP and provide the tractability analysis. Fourth, we propose an algorithmic framework based on Lagrangian duality and dynamic programming to facilitate problem solving. Unlike previous works, our approach contributes to delivering near-optimal solutions, as well as performance bounds on global optimum to demonstrate the quality of our near-optimal solutions. We use numerical results to illustrate the significant performance improvement of the proposed algorithm over existing NOMA and OFDMA schemes.

Our work extends previous study of user grouping in NOMA. In [18], the number of users to be multiplexed on a subcarrier is fixed, and performance evaluation consists of rule-based multiplexing policies. In our case, for each subcarrier, the number of users and their composition are both output from solving an optimization problem. Later in Section VII, results of optimized subcarrier assignment and user grouping will be presented for analysis. The current paper extends our previous study [12] in several dimensions. The extensions consist of the consideration of the WSR utility metric, a significant amount of additional theoretical analysis of problem tractability, the development of the performance bound on global optimality, as well as the consideration of user fairness in performance evaluation.

The rest of the paper is organized as follows. Section II gives the system models for single-carrier and multi-carrier NOMA. Section III formulates JPCAP for WSR utility and provides complexity analysis. Section IV analyzes the tractability for special cases of JPCAP. In Section V, we provide the tractability analysis for relaxations of JPCAP. Section VI proposes an algorithmic framework. Numerical results are given in Section VII. Conclusions are given in Section VIII.
Throughout the paper, we refer to subchannel interchangeably with subcarrier. We use $\mathcal{K}$ and $\mathcal{H}$ to denote the sets of users and subchannels, respectively, and $g_{kn}$ to denote the channel gain between the BS and user $k$ on subcarrier $n$. Let $p_{kn}$ be the power allocated to user $k$ on subcarrier $n$. A user $k$ is said to be multiplexed on a subchannel $n$, if and only if $p_{kn} > 0$. The power values are subject to optimization. At the receiver, each user equipment has MUD capabilities to perform multi-user signal decoding. With SIC, some of the co-channel interference will be treated as decodable signals instead of as additive noise.

### B. NOMA Systems

To ease the presentation of the system model, for the moment let us consider the case that all the $K$ users can multiplex on each subcarrier $n$ in a multi-carrier NOMA system (MC-NOMA) at downlink. For each subcarrier $n$, we sort the users in set $\mathcal{K}$ in the descending order of channel gains, and use bijection $b_n(k): \mathcal{K} \mapsto \{1, 2, \ldots, K\}$ to represent this order, where $b_n(k)$ is the position of user $k \in \mathcal{K}$ in the sorted sequence. For our downlink system scenario, in [25] (Chapter 6.2.2, pp. 238) it is shown that, with superposition coding, a user can decode the data of another user with worse channel gain, and this is not constrained by the specific power split. The reason is that the first user, due to its better receiving condition than the second user, can decode any data that the second user can successfully decode. Consider one subcarrier and two users $k$ and $h$ with gain $g_k > g_h$. User $h$ does not perform SIC, and its rate equals $\log(1 + \frac{p_k g_k}{p_h g_h + \eta})$, where $\eta$ denotes the noise, and $p_k$ and $p_h$ are the power levels. For user $k$ with better gain $g_k$, the SINR for the data for user $h$ is $\frac{p_k g_k}{p_h g_h + \eta} > \frac{p_k g_k}{p_h g_h + \eta}$. Hence user $k$ can decode the data (at the rate governed by the right-hand side of the inequality) of user $h$, and this is not dependent on the power relation. Thus user $k$ is able to perform SIC, by subtracting the re-encoded signal intended for receiver $h$ from the composite signal.

That is, user $k$ on subcarrier $n$, before decoding its signal of interest, first decodes the received interfering signals intended for the users $h \in \mathcal{K}\{k\}$ that appear later in the sequence than $k$, i.e., $b_n(h) > b_n(k)$. The interfering signals with order $b_n(h) < b_n(k)$ will not be decoded and thus treated as noise. Hence, the interference after SIC for user $k$ on subcarrier $n$ is $\sum_{h \in \mathcal{K}\{k\}; b_n(h) < b_n(k)} p_{hn} g_{hn} \forall k \in \mathcal{K}, \forall n \in \mathcal{H}$. If there are users having the same channel gain, then SIC applies following the principle in [25] (Chapter 6.2.2), provided that an ordering of the users is given. From the discussion, the SINR of user $k$ on subcarrier $n$ is given below.

$$\text{SINR}_{kn} = \frac{p_{kn} g_{kn}}{\sum_{h \in \mathcal{K}\{k\}; b_n(h) < b_n(k)} p_{hn} g_{hn} + \eta}$$  \hspace{1cm} (1)

The noise power here equals the product of the power spectral density of white Gaussian noise and the subcarrier bandwidth. The rate of each user in NOMA is determined by the user’s SINR after SIC. Thus, the achievable rate of user $k$ on subcarrier $n$ is $R_{kn} = \log(1 + \text{SINR}_{kn})$ nat/s with normalized bandwidth $\frac{B}{K} = 1$.

For single-carrier NOMA systems (SC-NOMA), we omit the subcarrier index. For convenience, the users $k \in \{1, \ldots, K\}$ in SC-NOMA are defined in the descending order of channel gains, where $g_1 \geq g_2 \geq \ldots \geq g_K$. Thus the user index also represents its position in the sequence, and user $k$ is able to decode the signal of user $h$ if $k < h$.

We define $\text{SINR}_k = \frac{p_k g_k}{\sum_{h \in \mathcal{K}\{k\}; h < k} p_h g_h + \eta}, \forall k \in \mathcal{K}$. The achievable rate of user $k$ is $R_k = \log(1 + \text{SINR}_k)$ nat/s with normalized bandwidth $B = 1$. Following the discussion earlier, for two users 1 and 2 with $g_1 > g_2$, the achievable rates are $\log(1 + \frac{p_1 g_1}{p_2 g_2 + \eta})$ and $\log(1 + \frac{p_2 g_2}{p_1 g_1 + \eta})$, respectively.

We use $u_k$ as a generic notation for the set of users multiplexed on subchannel $n$ for MC-NOMA. For SC-NOMA, the corresponding entity is denoted by $\mathcal{U}$. We use $M$, $1 \leq M \leq K$, to denote the maximum number of multiplexed users on a subcarrier. The reason of having this parameter is to address complexity considerations of implementing MUD and SIC. In NOMA, the system complexity increases by $M$, because a user device needs to decode up to $M$ signals. The setting of $M$ depends on receiver’s design complexity and signal processing delay for SIC [4], [8]. For practical implementation, $M$ is typically smaller than $K$. However, our optimization formulations and the solution algorithm are applicable to any value of $M$ between one and $K$. We also remark that the benefits of NOMA come with signaling overhead that is necessary to facilitate SIC. Although signaling is outside the scope of the current paper, it is of significance in practical implementation. On the other hand, it has been shown in [11] that NOMA remains superior to OMA in throughout when the signaling overhead is accounted for. In addition, parameter $M$, which limits the number of users per subcarrier, provides an effective way to control the signaling cost. By the results to be presented in Section VII, most of the improvements due to NOMA is achieved for small $M$.

Two utility functions, WSR and SR, are considered in this paper. The WSR utility is denoted by $f_W = \sum_{k \in \mathcal{K}} w_k \sum_{n \in \mathcal{N}} R_{kn}$, where $w_k$ is the weight coefficient of user $k \in \mathcal{K}$, Clearly, the selection of the weights has strong influence on the resource allocation among the users. In general, the weights can be used to steer the resource allocation towards various goals, such as to implement service class priority of users, and fairness (e.g., a user with averagely poor channel receives higher weight). In our work, the algorithmic approach is applicable without any assumption of the specific weight setting. For performance evaluation, we set the weights following proportional fairness. That is, for one time slot, a user’s weight is set to be the reciprocal of the average user rate prior to the current time slot [15]. As a result, the resource allocation will approach proportional fairness over time. The SR utility, a special case of WSR, is defined as $f_R = \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} R_{kn}$. The term of SR utility is used interchangeably throughout in this paper. For both SR and WSR, SIC with superposition coding [25] applies to the users multiplexed on the same subcarrier. As was discussed earlier, the decoding does not rely on assuming specific, a priori constraint on the power allocation among the users.
III. JOINT POWER AND CHANNEL ALLOCATION

In this section, we formulate JPCAP using WSR utility for MC-NOMA. We use W-JPCAP to denote the optimization problem. In general, JPCAP amounts to determining which users should be allocated to which subcarriers, as well as the optimal power allocation such that the total utility is maximized. In the following we define the variables and formulate W-JPCAP as $P1_{WSR}$ below, where all $p$-variables and $x$-variables are collected in vectors $p$ and $x$, respectively.

\[
\begin{align*}
pkn &= \text{allocated power to user } k \text{ on subcarrier } n. \\
xkn &= \begin{cases} 1 & \text{if user } k \text{ is multiplexed on subcarrier } n, \\
0 & \text{otherwise}. \end{cases}
\end{align*}
\]

\[
P1_{WSR}: \begin{array}{l}
\max_{x,p} \sum_{k \in A} \sum_{n \in C} R_{kn} x_{kn} \\
\text{s.t.} \sum_{k \in A} \sum_{n \in C} p_{kn} \leq P_{tot} \\
\sum_{n \in C} p_{kn} \leq P_k, \quad \forall k \in A \\
\sum_{k \in A} x_{kn} \leq M, \quad \forall n \in C
\end{array} \tag{2a}
\]

In $P1$, the objective (2a) is to maximize the WSR utility, where $R_{kn}$ is defined below.

\[
R_{kn} = \log(1 + \frac{p_{kn}g_{kn}}{\sum_{k \in A \setminus \{k\}} \frac{p_{kn}g_{kn}}{b_n(h) < b_h(k)}}), \quad \forall k \in A, \quad \forall n \in C \tag{3}
\]

Constraints (2b) and (2c) are respectively imposed to ensure that the total power budget and the individual power limit for each user are not exceeded. The per-user power limit $P_k$ is introduced for practical considerations, such as regulatory requirement on power towards a user device. Such a limit is also very common in OMA (e.g., [23], [24]). Constraints (2d) restrict the maximum number of multiplexed users on each subcarrier to $M$. We remark that the power allocation is represented by the $p$-variables of which the values are subject to optimization, whereas the power limits $P_{tot}$ and $P_k$, $k \in A$ are given entities. Suppose $M$ users, say users 1, ..., $M$, are allocated with positive power on channel $n$. If $P_{tot} \geq \sum_{k=1}^{M} P_k$ happens to hold, then all the $M$ users may be at their respective power limits, i.e., setting $p_{kn} = P_k$, $k = 1, \ldots, M$, is feasible. Otherwise, the $M$ users can still be allocated with positive power, though not all of them can be at the power limits. Indeed, if user $k$ is allocated with power $p_{kn} > 0$ on channel $n$, then typically $p_{kn} < P_k$ unless user $k$ is not allocated power on any other channel than $n$. For the total power limit $P_{tot}$ to be meaningful, one can assume $\sum_{k \in A} P_k > P_{tot}$ without loss of generality, because otherwise the total power limit $P_{tot}$ is not violated even if all users are allocated with their respective maximum power, that is, (2b) becomes void and should be dropped. We do not consider any further specific assumptions on the relation between $P_{tot}$ and $P_k, k \in A$, to keep the generality of the system model.

Remark 1: We do not explicitly impose the constraint that $p_{kn} > 0$ if and only if $x_{kn} = 1$. This is because setting $p_{kn} > 0$ and $x_{kn} = 0$ is clearly not optimal, by the facts that $p_{kn} > 0$ will lead to rate degradation of other users due to the co-channel interference (if there are other users on channel $n$), and that for user $k$, $p_{kn} > 0$ means power is consumed, but $x_{kn} = 0$ means no benefit as the rate in (2a) becomes zero.

Formulation $P1_{WSR}$ is non-linear and non-convex. The convexity of the objective function (2a) cannot be established in general, because of the presence of the binary $x$-variables and the product of $x$ and $p$. However, in complexity analysis, neither non-convexity nor non-linearity of a formulation proves the problem’s hardness, as a problem could be inappropriately formulated. Therefore, we provide formal hardness analysis for W-JPCAP below.

Theorem 2: W-JPCAP is NP-hard.

Proof: We establish the result in two steps. First, we conclude that if $M = 1$ in (2d), W-JPCAP is NP-hard, as it reduces to OFDMA subcarrier and power allocation, for which NP-hardness is provided in [22] and [23]. For general MC-NOMA with $M > 1$, we construct an instance of W-JPCAP and establish the equivalence between the instance and the OFDMA problem considered in [23]. We consider an instance of W-JPCAP with $K$ users, $N$ subcarriers, and $M = 2$. Let $\epsilon$ denote a small value with $0 < \epsilon < \frac{1}{KN}$. The total power $P_{tot}$ is set to $NK P_k$. The power limit $P_k$ is uniform for $\forall k \in A$, and the noise parameter $\eta = \epsilon$. Among the $K$ users, we select an arbitrary one, denoted by $\bar{k} \in A$, and assign a dominating weight $w_{\bar{k}} = e^{KN}$ and channel gain $g_{\bar{k}n} = 1$ on all the subcarriers, whereas the other users’ weights and channel gains are $w_k = \epsilon$ and $g_{kn} = 1$, $\forall k \in A \setminus \{\bar{k}\}$ and $\forall n \in C$. From above, the ratios $\frac{w_{\bar{k}}}{w_k}$ and $\frac{g_{\bar{k}n}}{g_{kn}}$ are sufficiently large such that allocating any power $p \leq P_{\bar{k}}$ to user $\bar{k}$ on any subcarrier $n$, the utility $w_{\bar{k}} R_{\bar{k}n} > \max(\sum_{k \in A \setminus \{\bar{k}\}} \sum_{n \in C} w_{kn} R_{kn})$ for using the same power budget $p$, since $\sum_{k \in A \setminus \{\bar{k}\}} \sum_{n \in C} w_{kn} R_{kn}$ is bounded by $KN e^{-KN} \log(1 + \frac{e^{KN}p}{\epsilon})$, and $w_{\bar{k}} R_{\bar{k}n} = e^{KN} \log(1 + \frac{e^{KN}p}{\epsilon})$ is clearly greater than $KN e^{-KN} \log(1 + \frac{e^{KN}p}{\epsilon})$. Thus, allocating power to user $\bar{k}$ rather than other users is preferable for maximizing utility.

Due to the uniform gain $g_{\bar{k}n}$ and the dominating weight $w_{\bar{k}}$, the optimal power allocation for user $\bar{k}$ is to uniformly allocate an amount of $P_{\bar{k}}$ to each subcarrier. Then the remaining problem is to allocate power $P_{tot} - P_{\bar{k}} = (NK - 1) P_k$ to the remaining $K - 1$ users. Every user $k \in A \setminus \{\bar{k}\}$ is still subject to constraint (2c). Note that for $M = 2$, each subcarrier now can accommodate one extra user at most. Compared to the OFDMA problem in [23], W-JPCAP has one extra total power constraint, i.e., (2b), however, recall that $P_{tot}$ is set to $NK P_k$, and for this value (2b) is in fact redundant. Therefore, a special case of W-JPCAP with $M > 1$ is equivalent to the OFDMA problem in [23], and the result follows.
IV. TRACTABILITY ANALYSIS FOR UNIFORM WEIGHTS

The hardness of W-JPCAP could have stemmed from several sources, e.g., the structure of the utility function, discrete variables, non-concave objective, and the constraints. We start from investigating how the weight in the utility function influences the problem’s tractability. The utility function can affect the computational complexity in problem solving [22], [24]. One example is that the SR maximization problem with total power constraint in OFDMA is polynomial-time solvable [22]. With WSR utility, solving the same problem is challenging [26]. In this section, we consider a special problem with total power constraint in OFDMA is polynomial-time solvable [22].

Intuitively, R-JPCAP appears somewhat easier than W-JPCAP, however, the tractability of R-JPCAP, for both SC-NOMA and MC-NOMA, is not known in the literature. In the following, analogously to W-JPCAP, we formulate R-JPCAP in $P_{1SR}$.

$$P_{1SR}: \max_{x_k, p_k} \sum_{k \in \mathcal{K}, n \in \mathcal{N}} R_{kn} x_{kn} \tag{4a}$$

s.t. (2b), (2c), (2d) \tag{4b}

First, we provide structural insights for power allocation for R-JPCAP in SC-NOMA.

**Theorem 3:** For R-JPCAP in SC-NOMA, the following hold at the optimum:

(a) Suppose $g_k \geq g_h$ for two users $k$ and $h$ with $k \neq h$, then the optimal power allocation satisfies $p_h > 0$ only if $p_k > 0$.

(b) Up to $M$ consecutive users in descending order of gain are allocated with positive power.

**Proof:** By the SC-NOMA system model, the SR utility function $f = f(p_1, \ldots, p_K)$ reads:

$$\log \left(1 + \frac{p_1 g_1}{\eta}\right) + \log \left(1 + \frac{p_2 g_2}{p_1 g_2 + \eta}\right) + \ldots + \log \left(1 + \frac{p_K g_K}{\sum_{h=1}^{K-1} p_h g_h + \eta}\right) = \log (p_1 g_1 + \eta) - \log \eta + \log ((p_1 + p_2) g_2 + \eta) - \log (p_1 g_2 + \eta) + \ldots + \log \left(\sum_{h=1}^{K-1} p_h g_h + \eta\right) - \log (p_1 g_1 + \eta)$$

Next, we consider the partial derivatives $\frac{\partial f}{\partial p_1}, \ldots, \frac{\partial f}{\partial p_K}$, as shown in (5).

$$\frac{\partial f}{\partial p_1} = \frac{g_1}{p_1 g_1 + \eta} - \frac{g_2}{p_1 g_2 + \eta} + \frac{g_3}{(p_1 + p_2) g_2 + \eta} - \frac{g_3}{(p_1 + p_2) g_3 + \eta}$$

For conclusion (a), since $g_1 \geq g_2 \geq \ldots \geq g_K$ for the single subcarrier in SC-NOMA, it can be easily checked from (5) that the partial derivatives $\frac{\partial f}{\partial p_1} \geq \frac{\partial f}{\partial p_2} \geq \frac{\partial f}{\partial p_3} \geq \ldots \geq \frac{\partial f}{\partial p_K} > 0$, irrespective of the power values. In general, from the partial derivatives, more utility will be obtained by increasing power $p_k$ instead of $p_h$ if $k < h$, $\forall k, h \in \mathcal{K}$. If the user with the best channel condition, i.e., user 1, has $p_1 \leq P_1$, the objective value $f(p_1, \ldots, p_K)_{\text{R}}$ can be improved by shifting power from other users to user 1 until $p_1$ equals the power limit $P_1$. Statement (a) also implies that the optimal power allocation will be in a consecutive manner, i.e., from user 1 to user $M$, one by one, and the result of (b) follows.

By Theorem 3, for the SR utility, the users being allocated positive power at optimum are consecutive in their gain values, starting from the user with the best channel gain. Note that the SR utility maximizes the throughput, but has the issue of fairness, which is addressed by the more general metric of weighted SR (WSR) utility, where the weights are set according to the proportional fairness policy. By Theorem 3 and its proof, R-JPCAP for SC-NOMA can be optimally solved by the procedure given in Algorithm 1. The users’ power allocation is performed in a consecutive manner, starting from user $k = 1$, and assigning power min($P_k, P_{\text{tot}}$) to

**Algorithm 1** Polynomial-Time Algorithm for R-JPCAP in SC-NOMA

1. Initialize $p_k^* = 0$, $\forall k \in \mathcal{K}$
2. for $k = 1 : M$
3. $p_k^* \leftarrow \min(P_k, P_{\text{tot}})$, $P_{\text{tot}} = P_{\text{tot}} - p_k^*$
4. if $P_{\text{tot}} = 0$ then
5. Break
6. Return: Optimal power allocation $p_1^*, \ldots, p_M^*$

$$f = \frac{g_{K-1}}{\sum_{h=1}^{K-1} p_h g_{K-1} + \eta} - \frac{g_K}{\sum_{h=1}^{K-1} p_h g_K + \eta} \geq 0$$

$$g_K \sum_{h=1}^{K} p_h g_K + \eta$$

$$\frac{\partial f}{\partial p_{K-1}} = \frac{g_{K-1}}{\sum_{h=1}^{K-1} p_h g_{K-1} + \eta} - \frac{g_K}{\sum_{h=1}^{K} p_h g_K + \eta} \geq 0$$

$$\frac{\partial f}{\partial p_K} = \frac{g_K}{\sum_{h=1}^{K} p_h g_K + \eta}$$
user $k$ and updating $P_{tot}$ accordingly. Algorithm 1 is clearly
of polynomial-time complexity, resolving the tractability of
R-JPCAP in SC-NOMA, giving Corollary 4 below.

Corollary 4: R-JPCAP for SC-NOMA is tractable, i.e.,
polynomial-time solvable.

Next, we analyze the computational complexity of R-JPCAP
in MC-NOMA.

Theorem 5: R-JPCAP for MC-NOMA is NP-hard.

Proof: The proof is analogous to that of Theorem 2.
For general R-JPCAP with $M > 1$, we construct a special
instance with $K$ users, $N$ subcarriers, $M = 2$, $P_{tot} = KN P_k$, and
uniform $P_k = 1$. We deploy a “dominant user” $\bar{k} \in \mathcal{K}$
in the instance, that is, user $\bar{k}$ has the highest and uniform
channel gain of $g_{k n} = 1$ for all the subcarriers, whereas
the channel gain of all the other users and subcarriers is
$g_{k n} \leq \frac{1}{P_k}$, $\forall k \in \mathcal{K}\setminus\{\bar{k}\}$, $\forall n \in \mathcal{N}$. One can observe that the
ratio of $\frac{g_{k n}}{g_{\bar{k} n}}$ for $\forall n \in \mathcal{N}$, $\forall k \in \mathcal{K}\setminus\{\bar{k}\}$ has been set sufficiently
large, and from (5), the partial derivative of user $\bar{k}$ satisfies
$\frac{\partial f}{\partial p_{\bar{k} n}} > \frac{\partial f}{\partial p_{\bar{k} n}}$ for $\forall k \in \mathcal{K}\setminus\{\bar{k}\}$ on each $n$, irrespective of
the power values. Then the statement (a) in Theorem 3 is
valid for any subcarrier $n$ in this instance, that is, if $\bar{k}$ is
multiplied on subcarrier $n$, the optimal power $P_{\bar{k} n} > 0$ only if
$p_{\bar{k} n} > 0$ for any $k \neq \bar{k}$. Thus, on each subcarrier $n$, allocating
power to user $\bar{k}$ is preferred for optimality. Furthermore, due
to the uniform channel gain for user $\bar{k}$, the optimal power
allocation for user $\bar{k}$ is to allocate equal power $\frac{P_{tot}}{N}$ on every
subcarrier. Then the remaining problem is equivalent to the
OFDMA resource allocation problem in [23], and the result
follows.

From the results in this section, using SR utility function
instead of WSR does not change the intractability of JPCAP.

V. TRACTABILITY ANALYSIS FOR RELAXED JPCAP

In this section, we aim to identify and characterize tractable
cases for JPCAP. We provide tractability and convexity analy-
sis for a relaxed version of W-JPCAP and R-JPCAP. For
problem’s relaxation, we make the following observations.
First, from the proofs of Theorem 2 and 5, we have the following corollary.

Corollary 6: Both R-JPCAP and W-JPCAP remain
NP-hard, even if constraint (2b) is relaxed (i.e., the constraint
is removed from the optimization formulations $P1_{SR}$
and $P1_{WSR}$).

Second, solving JPCAP will be challenging if (2c) is
present. The observation applies also to OFDMA resource
allocation, see e.g., [1], [22], [23]. Next, the discrete
$x$-variables are introduced in JPCAP due to the presence of
constraint (2d). This results in a non-convex feasible region.
In the following, we relax two constraints (2c) and (2d) as
well as removing $x$-variables, and construct a relaxed version of
R-JPCAP and W-JPCAP in $P2_{SR}$ and $P2_{WSR}$, respectively.

\begin{align}
P2_{SR} & : \max_{p} \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} R_{k n}, \quad \text{s.t. (2b)} \quad (6) \\
P2_{WSR} & : \max_{p} \sum_{k \in \mathcal{K}} w_k \sum_{n \in \mathcal{N}} R_{k n}, \quad \text{s.t. (2b)} \quad (7)
\end{align}

Note that both formulations above are with the $p$-variables
only. We characterize the optimal power allocation for $P2_{SR}$
in SC-NOMA first.

Lemma 7: For $P2_{SR}$ in SC-NOMA with $g_1 \geq g_2 \geq \ldots, \geq g_K$, power allocation $p_1 = P_{tot}$, $p_2 = \ldots = p_K = 0$ is
optimal.

Proof: First, observe that relaxing the user-individual
power constraint (2c) is equivalent to setting $P_k = P_{tot}$, i.e.,
the user power limit is set to be equal to the total power
limit, such that (2c) becomes redundant. Then, the result of
the lemma is obtained by applying the result of Theorem 3,
that is, the optimum can be computed by using Algorithm 1,
which, with $P_k = P_{tot}$, leads immediately to the result of the
lemma.

Next, we generalize the results of Lemma 7 to multi-carrier
systems, and show $P2_{SR}$ is also tractable in MC-NOMA.

Lemma 8: For $P2_{SR}$ in MC-NOMA, there is an optimal
solution satisfying $|\mathcal{U}_n| \leq 1$, $\forall n \in \mathcal{N}$, i.e., OMA is optimal.

Proof: Suppose at the global optimum, a subcarrier $n$
has $|\mathcal{U}_n| > 1$. Consider the two users having the largest gain
values in $\mathcal{U}_n$, and, without any loss of generality of the proof,
suppose that the two user indices are 1 and 2, with $g_{1n} \geq g_{2n}$.
Denote the power allocated to users 1 and 2 on this subcarrier
by $P_{1n}$ and $P_{2n}$, respectively, with $P_{1n} > 0$ and $P_{2n} > 0$.
Denote by $p^*_n$ the sum of the two, i.e., $p^*_n = P_{1n} + P_{2n}$.
The achieved sum utility for these two users is thus $R_n(2) = \log \frac{P_{1n} g_{1n} + \eta_{\max}}{\eta}$, whereas for user 2 the utility is zero. Comparing $R_n(1)$ and $R_n(2)$, we obtain the following.

\begin{align}
R_n(2) - R_n(1) & = \log(\frac{P_{1n} g_{1n} + \eta_{\max}}{p^*_n g_{2n} + \eta_{\max}}) - \log(\frac{P_{1n} g_{1n} + \eta_{\max}}{P_{1n} g_{2n} + \eta_{\max}}) \\
& = \log(\frac{P_{1n} g_{1n} p^*_n g_{2n} + P_{1n} g_{1n} \eta + p^*_n g_{2n} \eta + \eta^2}{P_{1n} g_{1n} P_{1n} g_{2n} + P_{1n} g_{1n} \eta + P_{1n} g_{2n} \eta + \eta^2}) \quad (8)
\end{align}

It can be observed that
\begin{align}
(p_{1n} g_{1n} + p^*_n g_{2n}) - (p^*_n g_{1n} + p_{1n} g_{2n})
& = (p_{1n} - p^*_n)(g_{1n} - g_{2n}) < 0 \quad (9)
\end{align}

Since $p_{1n} < p^*_n$ and $g_{1n} > g_{2n}$, we have $R_n(2) < R_n(1)$. This
contradicts the optimality of the first power allocation, and the
lemma follows.

From Lemma 8 for $P2_{SR}$, which amounts to maximizing
the sum rate utility subject to one single constraint on the
total power, OMA resource allocation is optimal. We remark
that $P2_{SR}$ is convex and tractable. It can be checked that the
Hessian matrix of the objective function in $P2_{SR}$ is negative
semi-definite. As the $p$-variables are continuous and (2b) is
linear, $P2_{SR}$ is convex. In general, the optimal solution can
be obtained by performing a polynomial-time algorithm [22],
that is, choosing the user with the best channel gain on
each subcarrier and then applying water-filling power allocation for the assigned users. The conclusion is summarized below.

Corollary 9: The optimization problem in $P_{2SR}$ is tractable for SC-NOMA and MC-NOMA.

Remark 10: In some proposed NOMA schemes, e.g., [9], [10], users with inferior channel condition may request more power to enhance fairness, e.g., if $g_{1n} \geq g_{2n} \geq \ldots \geq g_{Kn}$ on a subcarrier $n$, the power allocation is subject to $0 < p_{1n} \leq p_{2n} \leq \ldots \leq p_{Kn}$. By Lemma 7, we remark that, on each subcarrier $n$, equal power allocation for the multiplexed users $k \in \mathcal{U}_n$ is optimal.

In the following, we characterize the tractability and convexity for $P_{2WSR}$. From formulation $P_{2WSR}$, the convexity is not straightforward to obtain. Note that Lemma 7 and Lemma 8 may not hold for $P_{2WSR}$, as a user with inferior channel gain may be associated with higher weight, and as a result, the optimum possibly has $|t_{Kn}| > 1$ for some $n \in \mathcal{N}$. We make the following derivations to show $P_{2WSR}$ is convex.

Theorem 11: The optimization problem in $P_{2WSR}$ is convex.

Proof: From Eq. (3), for any subcarrier $n \in \mathcal{N}$, there are $K$ equations linking the power allocation with the interference term. For the user with the best gain value, there is no interference term in Eq. (3), and hence the power variable of this user can be expressed in its rate. Going through the remaining users in descending order of gain and performing successive variable substitution, the $p$-variables can be all expressed in the rate values. Utilizing the observation, we prove the convexity by reformulating $P_{2WSR}$ by treating rates $R_{kn}, k \in \mathcal{K}, n \in \mathcal{N}$ as the optimization variables. This transformation is analogous to the geometric programming method [27]. To facilitate the proof, we use $m_n(i)$ to denote the user index in the $i$th position in the sorted sequence for subcarrier $n$, with indices $i = 0, \ldots, K$, and the convention that $m_n(0) = 0$. Problem $P_{2WSR}$ is then reformulated below.

$$\max_R \sum_{i=1}^K \sum_{n=1}^N u_k \sum_{n=1}^N R_{kn}$$

s.t. $\sum_{n=1}^N \sum_{i=1}^K \left( \frac{\eta}{R_{mn}(i)} - \frac{\eta}{R_{mn}(i-1)} \right) \exp(\sum_{h=i}^K R_{mn}(h), n) \leq 0$ (10b)

$R_{kn} \geq 0, \forall k \in \mathcal{K}, \forall n \in \mathcal{N}$ (10c)

The objective and constraints (10c) are both linear. For constraints (10b), note that $\frac{\eta}{R_{mn}(i)} - \frac{\eta}{R_{mn}(i-1)} \geq 0$, due to the descending order of channel gains. Hence the sum-exponential function in (10b) is convex [28], and the theorem follows.

For the above convex and tractable cases, i.e., $P_{2SR}$ and $P_{2WSR}$, standard optimization approaches for convex problem can be applied. For intractable cases, we develop an algorithmic framework based on Lagrangian dual optimization and dynamic programming (DP) to provide both near-optimal solutions and optimality bounds in the next section.

VI. OPTIMIZATION ALGORITHM FOR NOMA POWER AND CHANNEL ALLOCATION

In view of the complexity results, we aim to develop an algorithm that is not for exact global optimum, yet the algorithm by design, is capable of providing near-optimal solutions. Moreover, the algorithm is expected to deliver optimality bounds in order to gauge performance, and is capable of progressively improving the bounds by scaling parameters.

In this section, we propose an algorithmic framework based on Lagrangian duality and dynamic programming (LDP). In the developed algorithm, we make use of the Lagrangian dual from relaxing the individual power constraint (2c) with multipliers, and we develop a DP based approach to solve the problem for given multipliers. The algorithm is designed to solve both R-JPCAP and W-JPCAP problems. For generality, we take W-JPCAP for illustration.

A. Lagrangian Duality and Power Discretization

Let vectors $p$ and $x$ collect all $p$-variables and $x$-variables, respectively. Vector $\lambda := \{\lambda_k, \forall k \in \mathcal{K}\}$ contains the Lagrangian multipliers associated with constraints (2c) in $P_{1WSR}$. We construct the subproblem of Lagrangian relaxation below.

$$P_{LR} : \max_{x, p} L(x, p, \lambda) = \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} u_k x_{kn} R_{kn}$$

$$+ \sum_{k \in \mathcal{K}} \lambda_k (P_k - \sum_{n \in \mathcal{N}} p_{kn})$$

s.t. (2b), (2d) (11)

$P_{LR}$ is subject to the total power constraint (2b) as well as constraints (2d) that limit the number of users in each subcarrier. Unlike the objective in $P_{1WSR}$ and $P_{1SR}$, allocating power to user $k$ on subcarrier $n$ in $P_{LR}$ requires to pay a penalty in utility, i.e., $-\lambda_k P_k$. The Lagrange dual function is defined by $z(\lambda) = \max L(x, p, \lambda)$. The dual optimum is correspondingly defined below.

$$z^* = \min_{\lambda \geq 0} z(\lambda)$$ (12)

The optimization task amounts to solving $P_{LR}$ for a given $\lambda$ and finding the optimal $\lambda$ to minimize the Lagrangian dual in (12). Note that since $P_{1WSR}$ and $P_{1SR}$ are non-convex in general, there may exist a duality gap between $z^*$ and global optimum $z^\dagger$ to the original problem, i.e., $z^* \geq z^\dagger$.

Formulation $P_{LR}$ is non-convex in general due to the reasons we discussed in Section IV. We consider solving $z(\lambda)$, making use of the observation that, once the power is discretized, $P_{LR}$ admits the use of DP for reaching optimality in polynomial-time of the problem size and the number of power discretization levels. To this end, we discretize the power budget $P_{tot}$ into $J$ uniform steps, and denote by $\delta$ the size of each step, i.e., $\delta = \frac{P_{tot}}{J}$. Denote by $p^l$ be the power value for level $j$ and $p^l = \delta \times j$, where $j \in J = \{1, \ldots, J\}$. We denote by $P_{LR,D}$ the version of $P_{LR}$ after power discretization. The formulation of $P_{LR,D}$ and its optimization variables
are presented below.

$$x_{kn}^j = \begin{cases} 1 & \text{if power level } j \text{ is allocated to user } k \text{ on subcarrier } n, \\ 0 & \text{otherwise.} \end{cases}$$

$$P_{LR-D} : \max_x L_D(x, \lambda) = \sum_{k \in \mathcal{K}} w_k \sum_{n \in \mathcal{N}} \sum_{j \in J} x_{kn}^j R_{kn}^j + \sum_{k \in \mathcal{K}} \lambda_k (P_k - \sum_{n \in \mathcal{N}} \sum_{j \in J} x_{kn}^j p_j)$$

subject to:

$$\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} x_{kn}^j p_j \leq P_{tot}$$

$$\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} x_{kn}^j \leq M, \forall n \in \mathcal{N}$$

$$\sum_{j \in J} x_{kn}^j \leq 1, \forall k \in \mathcal{K}, \forall n \in \mathcal{N}$$

The objective (13a) and constraints (13b), (13c) originate from $P_{LR}$ but are adapted to power discretization. In (13a), the achievable rate of allocating user $k$ on subcarrier $n$ with power level $j$ is denoted by $R_{kn}^j$ below for $\forall k \in \mathcal{K}, \forall n \in \mathcal{N}, \forall j \in J$, and with normalized bandwidth $\frac{B}{N} = 1$.

$$R_{kn}^j = \log(1 + \frac{p_j g_{kn}}{\sum_{h \in \mathcal{H}(k)} b_h(k) - b_n(k) \sum_{j' \in J} x_{kn}^{j'} p_j^{j'}}) g_{kn} + \eta$$

Constraints (13d) state that each user on a subcarrier can select one power level at most, and $\sum_{j \in J} x_{kn}^j = 0$ means that there is no power allocation for user $k$ on subcarrier $n$. For given $\lambda$, let $z_D(\lambda) = \max L_D(x, \lambda)$ and $P^*$ denote the optimal objective value and the corresponding power solution of $P_{LR-D}$, respectively. Next, we develop a DP based approach to solve $P_{LR-D}$ exactly to optimality.

**Remark 12:** Power discretization in $P_{LR-D}$ is considered as an approximation for the continuous power allocation in $P_{LR}$. However, in practical systems, the power is typically set in discrete steps, e.g., discrete power control in LTE downlink [29]. In this case the discrete model $P_{LR-D}$ is exact.

### B. Two-Stage DP Based Approach

Given power levels in set $J$ and multipliers in vector $\lambda$, problem $P_{LR-D}$ can be solved by using DP. In general, DP guarantees global optimality if the problem has the so-called “optimal substructure property” [30]. A classical example is the knapsack problem with integer coefficients [31]. In our case, $P_{LR-D}$ does exhibit the property, and a proof of the optimality of DP will be provided later in Theorem 13.

To ease the presentation, we describe the DP algorithm in two stages. In the first stage, intra-subcarrier power allocation is carried out among users, that is, for subcarrier $n \in \mathcal{N}$, the algorithm computes the optimal user power allocation by treating power $p_j^*$, $j = 1, \ldots, J$, as the power budget for the subcarrier in question. The optimal utility value of consuming power $p_j^*$ on $n$ is denoted by $V_{n,j}$, where $n \in \mathcal{N}$, $j \in J$. Since the number of multiplexed users cannot exceed $M$ on each subcarrier, we keep track on the optimum allocation for each $m \in \{1, \ldots, M\}$. We define a tuple of format $t = (u_t, M_t)$ to represent a candidate partial solution, where $u_t$ is the utility value, and $M_t$ is the number of users allocated with positive power. For a partial problem of assigning positive power to exactly $m$ out of the first $k$ users with power budget $p_j^*$, the optimal utility is denoted by $T_{k,j}^m$.

The optimality of stage 1 is obtained from DP recursion. The values $T_{k,j}^m$ can be arranged in form of a $K \times J \times M$ matrix $A_t$. Computing $T_{k,j}^m$ for $k, m, j = 1$ is straightforward. For $k \geq 2, m \geq 2$ and $j \geq 2$, the following recursive formula is used to obtain the corresponding value in $A_t$.

$$T_{k,j}^m = \max_{j' = 1, \ldots, j-1} \{ w_k R_{kn}^j - \lambda_k p_j^* + T_{k-1,j-j'}^{m-1} \}$$

From (14), the procedure of obtaining $T_{k,j}^m$ is decomposed into multiple stages, and the recursion is applied to move from one stage to another. Thus, each partial problem has an optimal substructure [30].

The algorithmic operations for stage one are given in Algorithm 2. The bulk of the computation starts at Line 2. For user $k = 1$, exactly one tuple is created for each $j$, and the corresponding utility value is $T_{1,j}^1$, $j \in J$, see Line 4. For users $k > 1$, the recursion is performed in Lines 5 to 15. In Lines 9 to 15, a tuple $t$ is created, and its utility value $u_t$ will replace the current $T_{k,j}^m$ if $u_t > T_{k,j}^m$. Note that in Line 2, the utility $u_t$ is the sum of two parts, i.e., assigning a trial power $p_j^*$ to user $k$ plus the previously obtained maximum utility $T_{k-1,j-j'}^{m-1}$. The algorithm terminates when all the $K$ users and $J$ power levels have been processed. The
Algorithm 3 Two-Stage Dynamic Programming (TSDP)

Input: $K, N, J, M$, and $\lambda$

Output: $z(\lambda)$

1: Initialize $\hat{T}_{n,j} \leftarrow 0$ for all $n \in \mathcal{N}, \forall j \in \{0, \ldots, J\}$

2: **Stage 1:**
3:   for $n = 1 : N$ do
4:     Perform Algorithm 2, and obtain $V_{n,1}, \ldots, V_{n,J}$
5:   end for

6: **Stage 2:**
7:   for $j = 1 : J$ do
8:     $\hat{T}_{1,j} \leftarrow V_{1,j}$
9:   end for
10: for $n = 2 : N$ do
11:     $\hat{T}_{n,j} = \max_{j'=1,\ldots,J} \{(V_{n,j'} + \hat{T}_{n-1,j-j'}), \hat{T}_{n-1,j}\}$
12:     $z_D(\lambda) \leftarrow \max_{n \in \mathcal{N}, j \in J} \hat{T}_{n,j}$

Algorithm 4 LDDP for Solving W-JPCAP

1: Initialize $\lambda$, tolerance $\epsilon$, number of iteration $C \leftarrow 0$, maximum number of iterations $C_{\text{max}}$, $\hat{d}$ and $\hat{d}$ such that $|\hat{d} - d| > \epsilon$, $p^* \leftarrow 0$
2: while $|\hat{d} - d| > \epsilon$ or $C \leq C_{\text{max}}$ do
3:     $\hat{d} \leftarrow d$
4:     Perform TSDP (Algorithm 3) to solve $P_{LR-D}$
5:     $z_D(\lambda) \leftarrow \max_{n \in \mathcal{N}, j \in J} \hat{T}_{n,j}$
6:     $p^* \leftarrow$ power solution of $z_D(\lambda)$
7:     $d \leftarrow z_D(\lambda)$
8: if $p^*$ violates constraints (2c) then
9:     Convert $p^*$ to $p_f$
10: Compute $f(p_f) = \sum_{k \in \mathcal{K}} w_k \sum_{n \in \mathcal{N}} R_{kn}$ by (3)
11: $V_{\text{UB}} \leftarrow f(p_f)$
12: Update $\lambda$ by subgradient method
13: $C \leftarrow C + 1$
14: end
15: Relax (13b) with $\mu$ and construct $P'_{LR-D}$
16: repeat
17:     Bisection search for $\mu$
18:     $z_D(\lambda, \mu) \leftarrow \max_{x} L_D'(x, \lambda, \mu)$
19: until $\mu^* \leftarrow \{\mu : \min_{x \geq 0} L_D'(x, \lambda, \mu)\}$
20: $V_{\text{UB}} \leftarrow z_D(\lambda, \mu^*)$
21: Return: $V_{\text{LB}}$ and $V_{\text{UB}}$

C. Algorithmic Framework: Lagrangian Duality

With Dynamic Programming

We develop a framework LDDP to deliver near-optimal solutions (N-LDDP) of the global optimum $z^*$ of $P_{1\text{WSR}}$ and $P_{1\text{SR}}$. We also derive a scheme UB-LDDP to provide upper bounds for gauging the solution quality of N-LDDP. LDDP is summarized in Algorithm 4. In Line 3 to Line 11, we obtain $z_D(\lambda)$ and the power solution $p^*$ by applying TSDP to solve $P_{LR-D}$. These steps constitute N-LDDP. The iterations for solving the Lagrangian dual terminate either after a specified number of iterations $C_{\text{max}}$, or if the difference between the objective values in two successive iterations is less than $\epsilon$ [30]. Lines 15–20 form UB-LDDP, which approximates the global optimum from above. UB-LDDP delivers an upper bound, i.e., a value that is guaranteed to be no smaller than the global optimum. The purpose is for performance evaluation. Note that the problem is maximization, and hence the solution from the first part of the algorithm, N-LDDP, has a utility value that is lower than global optimum. Computing the global optimum is NP-hard. However, to assess the performance of this solution, we can instead obtain the upper bound at significantly lower complexity. The deviation from global optimum cannot be more than the deviation from the upper bound, which can be used in our performance evaluation. 

In Line 9, the users’ individual power constraints may be violated in $p^*$. Thus, we develop a three-step approach to convert $p^*$ into a feasible solution $p_f$ for JPCAP if the former violates (2c). Let set $\mathcal{K}$ denote the users allocated with positive power in $p^*$. $\mathcal{K} \subseteq \mathcal{K}$. We denote $\mathcal{K}^* = \{k \in \mathcal{K} : \sum_{n \in \mathcal{N}} p_{kn} > P_k\}$ as the subset of users for which (2c) does not hold.

optimal value of assigning power $p_j$ on a subcarrier $n$ is stored in $V_{n,j}$. See Line 17.

In the second stage, power allocation of $P_o$ is carried out among the subcarriers, i.e., inter-subcarrier power allocation. Then DP is applied to perform optimal power allocation at the subcarrier level. For given $\lambda$, $z_D(\lambda)$ is obtained in Algorithm 3. The operations start at Line 3. In the first stage of TSDP, Algorithm 2 is performed to obtain $V_{n,1}, \ldots, V_{n,J}$ for each subcarrier $n \in \mathcal{N}$. Note that by the construction of Algorithm 2, in $V_{n,j}$, the number of multiplexed users is at most $M$ for all $n \in \mathcal{N}$, which is a stage in two, index $m$ is no longer needed. In Lines 6-10, based on the accumulated value $\hat{T}_{n,j}$, a new candidate for $T_{n,j}$ is obtained by adding $V_{n,j}$. Then for the partial problem of allocating power $p_j$ to 1, 2, ..., $n$ subcarriers, the optimal solution $\hat{T}_{n,j}$ is obtained.

The DP recursion (for $n \geq 2$) for the second stage is given in Line 10. The values $\hat{T}_{n,j}$ for all $k \in \mathcal{N}, j \in J$ can be viewed in form of an $N \times J$ matrix $\mathcal{A}_2$. From the DP recursions in TSDP, the global optimum of $P_{LR-D}$ is obtained from accumulating the solutions of the partial problems. By the end of stage two, $z_D(\lambda)$ is equal to the maximum $\hat{T}_{n,j}$ among the elements in $\mathcal{A}_2$, see Line 11. In $P_{LR-D}$, for the partial problem for users 1, 2, ..., $k$, subcarriers 1, 2, ..., $n$, and with a total power budget $p_j$, the optimum is independent of that for the remaining subcarriers or users. The complexity for computing optimality is provided in Theorem 13.

Theorem 13: The global optimum of $P_{LR-D}$ is obtained by TSDP with a time complexity being polynomial in $M, N, K, J$, and $J$.

Proof: The input of $P_{LR-D}$ are $N, K, M$ and $J$. By inspecting (14) and Line 10 in Algorithm 3, computing matrix $\mathcal{A}_1$ for all subcarriers requires $O(KNMJ^2)$ in running time. For matrix $\mathcal{A}_2$, the running time is of $O(NJ^2)$. Hence the former is dominating, and the overall time complexity is $O(KNMJ^2)$, which is polynomial in the input size.

Remark 14: Increasing $J$ provides better granularity in power discretization. By improving the granularity, the solution of $P_{LR-D}$ can approach arbitrarily close to that of $P_{LR}$. □

output discretization. By improving the granularity, the solution of $P_{LR-D}$ can approach arbitrarily close to that of $P_{LR}$.
To obtain $p_f$, step one, power allocation for each user $k \in \mathcal{K} \setminus \mathcal{K}'$, on subcarriers keeps same as in $p^*$. Step two, for each $k \in \mathcal{K}'$, the subcarriers are sorted in ascending order in power allocation of $k$. Following the sequence, power is allocated as in $p^*$, however until the limit $P_k$ is reached. Doing so releases an amount of $\sum_{k \in \mathcal{K}'} \sum_{n \in \mathcal{N}} P_k n - \sum_{k \in \mathcal{K}'} P_k$ power that can be re-allocated to users in $\mathcal{K} \setminus \mathcal{K}'$ in step three. The re-allocation follows the descending order of the product of channel gain and weight. That is, letting $(\tilde{k}, \tilde{n}) = \text{argmax} \ w_{kn} g_{kn}, k \in \mathcal{K} \setminus \mathcal{K}', n \in \mathcal{N}$, $p_{kn}$ is increased as much as allowed by $P_k$ and $P_{tot}$, then we select the next best candidate, and so on, until either $P_{tot}$ or $P_k$ for all $k \in \mathcal{K}$ is reached. In Line 10, we obtain the resulting utility $f(p_f \tilde{w})$ for using $p_f$, where the calculation of $f(p_f \tilde{w}) = \sum_{k \in \mathcal{K}} w_k \sum_{n \in \mathcal{N}} R_{kn}$ follows the equation in (3). The utility value of N-LDDP is delivered in $V_{LB}$ at Line 11. In Line 12, $V_{LB}$ is used in the calculation of step size of subgradient optimization (see [30]).

To provide an upper bound for $z^\dagger$, we apply post-processing to $z_D(\lambda)$ in LDDP. We remark that if $J$ is sufficiently large, $z_D(\lambda)$ can be empirically considered as an upper bound to W-JPCAP or R-JPCAP due to Lagrangian duality, however, theoretically there is no guarantee. For example, $z_D(\lambda)$ could be possibly less than $z^\dagger$ for small $J$, e.g., $J = 1$. From Line 15 to Line 20, we design an approach to convert $z_D(\lambda)$ to a theoretically guaranteed upper bound. First, in Line 15, for the multipliers in $\lambda$ and keeping their values fixed, we further relax the total power constraint (13b) with multiplier $\mu$, and reconstruct the subproblem of Lagrangian relaxation below.

$$P_{LR-D}' : \max_{x, \lambda, \mu} L_D'(x, \lambda, \mu)$$

$$= \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} \sum_{j \in J} R_{kn} \tilde{R}_{kn}^j$$

$$+ \sum_{k \in \mathcal{K}} \lambda_k (P_k - \sum_{n \in \mathcal{N}} \sum_{j \in J} R_{kn} \tilde{R}_{kn}^j)$$

$$+ \mu (P_{tot} - \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} \sum_{j \in J} R_{kn} \tilde{R}_{kn}^j)$$

s.t. (13c) and (13d)

The calculation of $\tilde{R}_{kn}^j$ for $\forall k \in \mathcal{K}, \forall n \in \mathcal{N}$, and $\forall j \in J$ is shown below.

$$\tilde{R}_{kn}^j = (1 + \frac{(p^j + \delta) g_{kn}}{\sum_{k' \in \mathcal{K} \setminus \{k\}} \sum_{j' \in J} \sum_{n' \in \mathcal{N}} \sum_{j' \in J} (g_{kn}^j p^j - \delta g_{kn}^j)) g_{kn} + \eta)}$$

Recall that $\delta$ is the step in power discretization. In comparison to $L(x, p, \lambda)$ in $P_{LR}$ and $L_D(x, \lambda)$ in $P_{LR-D}$, the construction of $P_{LR-D}'$ contains addition or subtraction of power $\delta$, for the purpose of ensuring that the outcome is a valid upper bound to the global optimum. This is achieved by using $\delta$ to obtain more optimistic values in all three parts of the function. In (16), for example, one power step $\delta$ is added to the signal of interest in the numerator and each interfering signal becomes weaker due to the subtraction of $\delta$ in the denominator. As a result, the overall utility for $L_{D}'(x, \lambda, \mu)$ is guaranteed to be an over-estimation. From Lines 16 to 19, bisection search is applied to obtain the optimal $\mu^*$ such that $z_D(\lambda, \mu^*) = \min \mu \geq 0 X_{UB}$ in Line 20. The validity of this upper bound is proved below.

**Theorem 15:** $V_{UB} \geq z^\dagger$.

**Proof:** Suppose $\lambda^*$ is the multiplier vector for $P_{LR-D}$ when Algorithm 4 terminates. Note that $\lambda^*$ for $z(\lambda^*)$ may not necessarily lead to the minimum dual value $z^\dagger$ in $P_{LR}$, so we have $z(\lambda^*) \geq z^\dagger$. We prove that $z_D(\lambda^*, \mu^*) \geq z(\lambda^*)$ holds, to show $V_{UB}$ is an upper bound of $z^\dagger$. We use vector $p_c > 0$ to denote the optimal power allocation for $z(\lambda^*)$. Based on $p_c$, we now construct a power vector $p_d > 0$ for $P_{LR-D}'$. Each power value in $p_d$ is rounded to $p_d$ such that each element in $p_d$ is represented by the closest power level $j \in J$, i.e., $p_d + \theta = p_d$, where $|\theta| \leq \delta$. Given $p_c$ and $p_d$, the corresponding x-vectors $x_c$ and $x_d$ are derived for $P_{LR}$ and $P_{LR-D}'$. Letting $\lambda^*$ and $\mu^*$ $x_d$ in the objective of $P_{LR-D}'$, we have $L_d'(x_d, \lambda^*, \mu)$. Since we over-calculate the objective in $P_{LR-D}'$, the summation of part I and part II in $L_d'(x_d, \lambda^*, \mu)$ is greater than or equal to $L(x_c, p_c, \lambda^*)$. Part III is no less than zero in $L_d'(x_d, \lambda^*, \mu)$ for any $\mu \geq 0$ due to $\| p_c \|_1 \leq P_{tot}$ and $|\theta| \leq \delta$. Then $L_d'(x_d, \lambda^*, \mu) \geq L(x_c, p_c, \lambda^*) = z(\lambda^*)$ for any $\mu \geq 0$. Thus, we have $L_d'(x_d, \lambda^*, \mu^*) \geq z(\lambda^*)$. Note that for $\lambda^*$ and $\mu^*$, $V_{UB} = z_D(\lambda^*, \mu^*)$ is the optimum value of the Lagrangian relaxation in $P_{LR-D}'$, whereas $L_d'(x_d, \lambda^*, \mu^*)$ is not because $x_d$ is not necessarily an optimal power allocation. Therefore, $z_D(\lambda^*, \mu^*) \geq L_d'(x_d, \lambda^*, \mu^*)$, and the conclusion follows.

On the complexity of LDDP, we observe the following. Within each iteration, N-LDDP calls Algorithm TSDDP that has polynomial-time complexity $O(K NM J^2)$ and hence is scalable. Note that the number of power levels $J$ can be tuned from the complexity perspective. An iteration of N-LDDP may require the conversion to feasible power allocation (Line 9). It is easily observed that this three-step conversion, as outlined earlier, has a complexity of $O(K N \log_2(K N))$, which scales much better than $O(K NM J^2)$. Next, obtaining the upper bound in UB-LDDP consists of using Algorithm TSDDP in one-dimensional bisection section. This computation does not lead to the computational bottleneck, because the upper bound is computed only once, and its purpose is not for power allocation in NOMA, but for performance evaluation as a post-processing step. Hence, overall, the complexity is determined by N-LDDP, and equals $O(K N M J^2)$, where $C$ is the number of subgradient optimization iterations upon termination. Subgradient optimization for Lagrangian duality has asymptotic convergence in general. In the next section, however, we observe that convergence is approached with only a few iterations.

VII. PERFORMANCE EVALUATION

A. Experimental Setup

We have carried out performance studies in downlink with randomly and uniformly distributed users. Table I summarizes
the key parameters. We generate one hundred instances and consider the average performance. To evaluate the performance of LDDP, we have implemented a previous NOMA power and channel allocation scheme called “fractional transmit power control” (NOMA-FTPC) and an OFDMA scheme with FTPC (OFDMA-FTPC) [10]. In these two schemes, the set of multiplexed users $U_n$ for each subcarrier $n$ is determined by a greedy-based user grouping strategy, where $|U_n| = M$ for NOMA-FTPC and $|U_n| = 1$ for OFDMA-FTPC. Based on the user allocation, the FTPC method is then used for power allocation. In FTPC, more power is allocated to the users with inferior channel condition for the fairness consideration [10].

For OFDMA-FTPC, following the LTE standard, the overall bandwidth of 4.5 MHz is divided into twenty-five subchannels with the bandwidth of 180 kHz for each. For NOMA implementation, considering the fact that the decoding complexity and signaling overhead increase with the number of subcarriers [6], and following the NOMA setup in [10], we consider five subcarriers with the bandwidth of 900 kHz for each in NOMA-FTPC and in the proposed LDDP.

In the simulations, N-LDDP aims to deliver a near-optimal solution (also a lower bound), and UB-LDDP, by design, provides an upper bound for global optimum. In the following, we examine five performance aspects. First, a comparative study for the SR utility of LDDP, NOMA-FTPC and OFDMA-FTPC is carried out. Second, we evaluate the convergence behavior of LDDP. Third, we consider the WSR utility and examine users’ fairness. Fourth, we evaluate the throughput performance for the cell-edge users. Fifth, we investigate the characteristics of user grouping in N-LDDP.

### B. Performance in Throughput and Bounding

Applying the three algorithms to all instance, the average results are summarized in Fig. 1 to 3. In Fig. 1, we evaluate the SR utility with respect to user number $K$, with setting $M = 2$ and $J = 100$. We make the following observations. First, the performance improvement tends to be marginal for larger $K$ in all the schemes. This is expected since the multiuser diversity is effective when the number of users is small, and is saturated if $K$ is large. Second, N-LDDP outperforms NOMA-FTPC and OFDMA-FTPC. N-LDDP achieves performance improvement of around 20% over NOMA-FTPC. NOMA-FTPC, in turn, performs much better than OFDMA. Third, N-LDDP is capable of providing near-optimal solutions. The average gap between UB-LDDP and N-LDDP is 11% in average, and the variation of the gap is insensitive to the number of users. This implies that the gap between N-LDDP and the global optimum $z^*$ is even smaller than 11%.

Next, the impact of parameters $J$ and $M$ is evaluated. The instances with $K = 20$ are used in the simulations, and $M = 2$ in Fig. 2 and $J = 100$ in Fig. 3. From Fig. 2, increasing $J$ leads to progressively tighter intervals between UB-LDDP and N-LDDP since larger $J$ provides better granularity in power discretization, and thus improves the solution quality. Moreover, we observe that the improvement comes mainly from UB-LDDP. This is because, for delivering the upper bound, the objective value in $P_{UB,D}$ has been intentionally over-calculated by $\delta = \frac{P_{tot}}{J}$. Compared to using $\delta = 0$, applying $\delta > 0$ in the over-calculation results in an excess of utility in the objective. This excess part is clearly $J$-related, and is significantly reduced when $J$ is large.

In Fig. 3, more users are allowed to share the same subcarrier. Increasing $M$ leads to more total throughput for both NOMA schemes. One can observe that N-LDDP constantly

### Table I

**Simulation Parameters**

| Parameter               | Value                  |
|-------------------------|------------------------|
| Cell radius             | 200 m                  |
| Carrier frequency       | 2 GHz                  |
| Total bandwidth ($B$)   | 4.5 MHz                |
| Number of subcarriers ($N$) | 5 in NOMA, 25 in OFDMA |
| Number of users ($K$)   | 4 to 20                |
| Path loss               | COST 231-HATA          |
| Shadowing               | Log-normal, 8 dB standard deviation |
| Fading                  | Rayleigh flat fading [25] |
| Noise power spectral density | -173 dBm/Hz           |
| Total power ($P_{tot}$) | 1 W                    |
| Number of power levels ($J$) | 20 to 100              |
| Minimum power unit ($\delta = \frac{P_{tot}}{J}$) | 0.01 to 0.05 W         |
| User power limit ($P_k$) | 0.2 W                  |
| Parameter $M$           | 2 to 6                 |
| Tolerance $\epsilon$ in LDDP | $10^{-5}$             |
| $C_{max}$ in LDDP       | 200                    |

![Fig. 1. LDDP performance in respect of the number of users.](image1)

![Fig. 2. LDDP bounding performance in respect of parameter J.](image2)
outperforms NOMA-FTPC. We also notice that increasing $M$ results in degradation of UB-LDDP. The main reason is that when $M$ grows, the over-calculations in the objective of $P_{LR-D}^*$ are accumulated over all multiplexed users.

C. Performance in Convergence

We illustrate the convergence behavior of LDDP by two representative instances with 4 and 20 users, respectively, with $N = 5$, $J = 100$, and $M = 2$. The evolution of the values of $z_D(\lambda)$ and $V_{LB}$ over the iteration number $C$ is provided in Fig. 4.

From the figure, a majority of the iterations is part of the tailing-off effect. The utility ($V_{LB}$ of the algorithm) and the Lagrangian dual function ($z_D(\lambda)$) both approach the achievable values with 10 iterations or fewer. Note that each iteration has polynomial-time complexity, see Theorem 13.

D. Performance in Fairness

As our next part of results, we evaluate the performance for LDDP, NOMA-FTPC, and OFDMA-FTPC in fairness. We examine the fairness over a scheduling time period. As was mentioned earlier, scheduling in the time domain is slotted. Denote by $t$ the time slot index. We define a scheduling frame consisting of 20 slots. The channel state information is collected once per frame. Define $\bar{R}_k(t) = (1 - \frac{1}{T})\bar{R}_k(t - 1) + \frac{1}{T}r_k(t - 1)$ as user $k$’s average rate prior to slot $t$, where parameter $T$ is the length of a time window (in the number of time slots), and $r_k(t - 1)$ is user $k$’s instantaneous rate in slot $t - 1$ [15]. By proportional fairness, the weight of user $k$ in slot $t$ is set to $1/\bar{R}_k(t)$, and for each slot, N-LDDP, NOMA-FTPC, and OFDMA-FTPC with WSR utility maximization are performed once.

Suppose the average users’ rates are $\bar{R}_1, \ldots, \bar{R}_K$ at the end of the scheduling period, and consider the Jain’s fairness index, computed as $\frac{\sum_{k=1}^{K} \bar{R}_k^2}{K \sum_{k=1}^{K} \bar{R}_k^2}$. This index, developed in [32], is widely used as a fairness measure for user throughput in communications networks. The value of this fairness index is between $\frac{1}{K}$ and 1.0. A higher value indicates fairer throughput distribution, and the maximum value of 1.0 is reached if and only if all users achieve exactly the same throughput. Note that the use of this index, by itself, does not prevent a user from being served with low throughput (or even zero throughput), which, however, will most likely bring down the value of the index. In our case, zero or very low throughput of any user is avoided by the fact that the weights in W-JPCAP are set in accordance with proportionally fair scheduling. Hence, over time, the weight of a user will increase to infinity, if the user keeps being allocated with zero throughput. For further insights of user throughput, particularly throughput of cell-edge users, please see Section VII-E.

The fairness index in respect of $K$ and $M$ is shown in Fig. 5 and Fig. 6, respectively. The parameters are set to be $M = 2$ in Fig. 5 and $K = 20$ in Fig. 6. For both figures, we consider a scheduling period of 100 time slots, with $J = 100$ and $T = 50$. From Fig. 5 and 6, we observe that, first, the proposed N-LDDP achieves the best performance. Moreover, Fig. 5, increasing $K$ leads to fairness degradation in all schemes due to more competition among the users.
From Fig. 6, the fairness index increases in $M$. This is because a larger $M$ provides more flexibility in resource allocation among the users. There is however a saturation effect, showing that the constraining impact due to limiting the number of multiplexed users per subcarrier decreases when the limit becomes large. Note that in Fig. 3, the improvement of throughput is marginal with $M$ becomes large. Thus a moderate $M$ is justified not only by implementation complexity, but also that having a large $M$ may not lead to significant performance improvement.

In the two figures, OFDMA-FTPC gives the lowest fairness index. A particular reason is that the FTPC channel and power allocation scheme is sub-optimal. This also explains the improvement enabled by the proposed power optimization algorithm in comparison to NOMA-FTPC. We also remark that the vertical axis of the two figures starts from a positive value, hence the relative difference between the schemes in fairness is smaller than what it may appear to be.

E. Performance for Cell-Edge Users in Throughput

The performance of cell-edge users is of significance. To evaluate, we split the service area of the cell into an edge zone and a center zone. Performance comparison is carried out for twenty-user instances with $M = 2$, $J = 100$, and 100 time slots. In the simulations, we deploy half of the users to the cell edge in each instance. Each value in Fig. 7 represents the average user rate over the entire scheduling period. We observe that using LDDP significantly improves the rates for cell-edge users. From the results in Fig. 7, the average rates of all cell-edge users in N-LDDP are much more than those of NOMA-FTPC and OFDMA-FTPC.

F. Characteristics of User Grouping

The final part of our performance study is on characteristics of user grouping, i.e., which users tend to be multiplexed together on the same subcarrier by optimization. We consider a scenario with $M = 2$, $J = 100$, $K = 20$, and apply N-LDDP to 1,000 realizations. For each subcarrier, we index the users in descending order of channel gain. If there are two users multiplexed on the subcarrier in the algorithm solution, we consider the difference of the two user indices. For example, if the users with the highest and lowest gains are grouped together, the difference is 19. The results are illustrated in Fig. 8, where the horizontal axis is the difference in index.

From the figure, we observe that users having large difference in channel gain are more likely to be multiplexed on the same subcarrier. This is coherent with the conclusion in [18]. On the other hand, it is also evident by the figure that optimal assignment is not necessarily to select users with the best and poorest gains on a subcarrier. The observation motivates the treatment of subcarrier allocation as optimization variables.

VIII. CONCLUSIONS

We have considered jointly optimizing power and channel allocation for NOMA. Theoretical insights on complexity and optimality have been provided, and we have proposed an algorithm framework based on Lagrangian dual optimization and dynamic programming. The proposed algorithm is capable of providing near-optimal solutions as well as bounding the global optimum tightly. Numerical results demonstrate that the proposed algorithmic notions result in significant improvement of throughput and fairness in comparison to existing OFDMA and NOMA schemes.

An extension of the work is the consideration of max-min fairness for one scheduling instance. In this case, one solution approach is to perform a bi-section search. For each target level that represents the minimum throughput required for all users, the problem reduces to a feasibility test, i.e., whether or not the target is achievable subject to the power limits. This can be formulated as to minimize the total power, with constraints specifying the throughput target value and user-individual power limits. The development of optimization algorithms for this problem is subject to further study.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their comments and suggestions.

REFERENCES

[1] D. Yuan, J. Joung, C. K. Ho, and S. Sun, “On tractability aspects of optimal resource allocation in OFDMA systems,” IEEE Trans. Veh. Technol., vol. 62, no. 2, pp. 863–873, Feb. 2013.
[2] L. Lei, D. Yuan, C. K. Ho, and S. Sun, “A unified graph labeling algorithm for consecutive-block channel allocation in SC-FDMA,” IEEE Trans. Wireless Commun., vol. 12, no. 11, pp. 5767–5779, Nov. 2013.
Chin Keong Ho (S’05–M’07) received the B.Eng. degree (Hons.) with a minor in business administration and the M.Eng. degree from the Department of Electrical Engineering, National University of Singapore, in 1999 and 2001, respectively, and the Ph.D. degree from the Eindhoven University of Technology, The Netherlands, in 2009, where he concurrently conducts research work with Philips Research. Since 2000, he has been with the Institute for Infocomm Research, Agency for Science, Technology, and Research, Singapore. His research interests include green wireless communications with focus on energy-efficient solutions and with energy harvesting constraints, cooperative and adaptive wireless communications, implementation aspects of multicarrier and multi-antenna communications, and innovative use of communications in the field of robotics. His works received the Best Student Journal Paper Award of the IEEE Sweden VT-COM-IT Joint Chapter in 2014, the IEEE Marconi Prize Paper Award in 2015, and the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2016.

Sumei Sun (SM’12–F’16) is currently the Head of the Communications and Networks Cluster, Institute for Infocomm Research, Agency for Science, Technology, and Research, Singapore. Her research focus is on energy- and spectrum-efficient communication technologies for connecting human, machines, and things. She is an inventor and a co-inventor of 30 granted patents and over 30 pending patent applications, many of which have been licensed to industry. She has authored and co-authored over 200 technical papers in prestigious IEEE journals and conferences. She has also been actively contributing to organizing conferences in different roles. Some of her recent conference services include the Executive Vice Chair of Globecom 2017, Symposium Co-Chair of ICC 2015 and 2016, Track Co-Chair of the IEEE VTC 2012 Spring, VTC 2014 Spring and VTC 2016 Fall, and Publicity Co-Chair of PIMRC 2015. She is an Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (TVT) since 2011, an Editor for the IEEE WIRELESS COMMUNICATION LETTERS from 2011 to 2016, and an Editor of the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS since 2015. She received the Top Editor Award in 2016, Top15 Outstanding Editors recognition in 2014, and Top Associate Editor recognition in 2013 and 2012, all from TVT. She is a Distinguished Lecturer of the IEEE Vehicular Technology Society 2014–2018, a co-recipient of the 16th PIMRC Best Paper Award, and a Distinguished Visiting Fellow of the Royal Academy of Engineering, U.K., in 2014.