Enhanced Uplink Resource Allocation in Non-Orthogonal Multiple Access Systems

Rukhsana Ruby, Member, IEEE, Shuxin Zhong, Hailiang Yang, and Kaishun Wu, Member, IEEE

Abstract—Non-orthogonal multiple access (NOMA) is envisioned to be one of the most beneficial technologies for next generation wireless networks due to its enhanced performance compared to other conventional radio access techniques. Although the principle of NOMA allows multiple users to use the same frequency resource, due to decoding complication, information of users in practical systems cannot be decoded successfully if many of them use the same channel. Consequently, assigned spectrum of a system needs to be split into multiple subchannels in order to multiplex that among many users. Uplink resource allocation for such systems is more complicated compared to the downlink ones due to the individual users’ power constraints and discrete nature of subchannel assignment. In this paper, we propose an uplink subchannel and power allocation scheme for such systems. Due to the NP-hard and non-convex nature of the problem, the complete solution, that optimizes both subchannel assignment and power allocation jointly, is intractable. Consequently, we solve the problem in two steps. First, based on the assumption that the maximal power level of a user is subdivided equally among its allocated subchannels, we apply many-to-many matching model to solve the subchannel-user mapping problem. Then, in order to enhance the performance of the system further, we apply iterative water-filling and geometric programming two power allocation techniques to allocate power in each allocated subchannel-user slot optimally. Extensive simulation has been conducted to verify the effectiveness of the proposed scheme. The results demonstrate that the proposed scheme always outperforms all existing works in this context under all possible scenarios.

Index Terms—NOMA Systems; Optimal Resource Allocation; Many-to-Many Matching Model; Geometric Programming; Iterative Water-Filling Algorithm

I. INTRODUCTION

The data traffic over cellular networks is projected to grow explosively in the coming years due to the proliferation of smartphones, tablets, smart terminals and emerging applications (e.g., machine-type-communications (MTC)) [11–13]. Consequently, future radio access networks [5, 6] are expected to have the capability of supporting massive connectivity, diverse sets of users and applications with radically different requirements in terms of delay, bandwidth and so on. In order to obtain fruitful outcome in this context, designing an effective and efficient radio access technology [7] is one of the possible solutions. Through experimental and theoretical analysis, it is proved that non-orthogonal multiple access (NOMA) technique is able to provide enhanced performance comparing with other orthogonal multiple access (OMA) techniques, such as time division multiple access (TDMA) and frequency division multiple access (FDMA) [8–11]. Consequently, NOMA is considered as the future dominating radio access technique, and is expected to satisfy the ever-increasing demands of future cellular networks.

Conceptually, power-domain NOMA allows multiple users to occupy the same frequency channel. By applying successive interference cancellation (SIC) [12] in NOMA systems, superposition coded signal can be correctly decoded and demodulated at the receiver. Although NOMA technique allows multiple users to be superimposed on the same frequency channel, due to the error propagation in the SIC technique, it is not an optimal design to assign large number of users on the same channel. Consequently, dedicated spectrum of a system needs to be subdivided into multiple subchannels in order to support increased number of users. At the same time, how to allocate these subchannels among users in a multiplex manner given the allowable maximum number of users that can use a subchannel simultaneously, is an important problem. Extensive research has been conducted on the downlink subchannel and power allocation for such NOMA systems. Based on some assumption of having constant power on the subchannels, typically, existing works provide some heuristics for subchannel-user mapping task. Once the subchannel-user mapping information is known, in order to enhance the performance of the system further, different existing works have provided different schemes for power allocation across the allocated subchannel-user slots. For example, in [13–14], the authors use fractional transmit power allocation technique among users and equal power allocation concept across subchannels. [15] uses water-filling-based approach for power allocation, and in [16], the authors use difference of convex (DC) programming-based [17] approach for the power allocation in both user and subchannel levels. Energy-efficient downlink resource allocation has also been studied in some papers, such as [18–20].

Unlike the downlink one, uplink resource allocation even in conventional OMA systems is considered challenging [21–23] because of the individual users’ power constraints and discrete nature of subchannel assignment. On the other hand, decoding technique in NOMA systems, SIC, is a multi-user detection technique that uses the structured nature of interference to decode multiple concurrent transmissions. Each individual signal from the composite signal is retrieved one by one following some order. If any of the signals is failed

This research was supported in part by the China NSFC Grant 61472259, Guangdong Natural Science Foundation (No. 2017A030312008, 2016A030313036), Shenzhen Science and Technology Foundation (No. JCYJ201703021404946299, JCYJ2017041210753954, JCYJ20150324140036842), Guangdong Talent Project 2014TQ01XZ38, 2015TX01X111, GDUPE(2015). Kaishun Wu is the corresponding author. R. Ruby, S. Zhong, H. Yang and K. Wu are with the college of Computer Science and Software Engineering, Shenzhen University, Shenzhen, Guangdong, 518060 China, e-mail: ruby@szu.edu.cn
to be decoded, it is unlikely that the rest of the signals can be decoded. Therefore, decoding order plays the significant role on the success of decoding operation and the throughput achieved by each individual signal. All these complications bring further difficulties in the uplink resource allocation of NOMA systems compared to OMA systems. Although NOMA principle does not enforce the decoding order of received superposition coded signals, it is proved in [24] that decoding of stronger signals ahead of weaker signals is beneficial for the system in terms of throughput and proportional fairness.

Compared to the downlink resource allocation in NOMA systems, uplink resource allocation is not that much studied. Still, there are some works in this context. Unlike the system in our work (in which NOMA technique is employed in the frequency domain), for a system in which NOMA technique is employed in the time domain, a set of uplink resource allocation schemes is provided in [24], [25] with the objective of throughput maximization and fairness of the system. Although the scheduling scheme in [25] assumes that the system has one time slot and a set of users with their power constraints to schedule, the scheduling scheme in [24] optimizes the total throughput and fairness of the system over a set of time slots and users. On the other hand, the main drawback of the work in [24] is, the time slots are resource elements and are invariant over time, which is very impractical for wireless networks. Moreover, given the power constraint of each individual user, each user can get only one time slot (i.e., one resource block), the concept of which fails to exploit multi-user-channel diversity of wireless systems. However, in practice, if multiple time slots are allocated to a user, the performance of that user may be enhanced. Another very close work compared to our work is [26]. In this work, the authors have proposed an uplink subchannel and power allocation scheme based on iterative water-filling technique [27]. With the expectation of utilizing the multi-user-channel diversity, this resource allocation scheme overcomes the drawback of the solution in [24] by assigning multiple subchannels (i.e., multiple resource blocks) to each user. Moreover, this scheme assigns exactly the maximum allowable number of users to each subchannel and gives more preference to the users with better channel while solving the subchannel-user mapping problem. However, even in uplink OMA systems [22], we previously observed that not necessarily the more the users allocated to each subchannel, the better the throughput is, especially in worse channel condition. This is because each user requires to subdivide its limited power level among its allocated subchannels. Furthermore, in NOMA systems, giving less privilege to the users with worse channel not necessarily enhances the throughput. Since the power level of other users in such systems is considered as interference for some particular user, pairing users with highly different channel condition is sometimes conducive to the performance of the system. In this paper, our objective is to overcome the drawbacks of existing aforementioned uplink resource allocation schemes, and to take NOMA-specific all useful scheduling insights into account.

The contribution of this paper is an elegant uplink subchannel and power allocation scheme in a NOMA system with enhanced performance. Since this problem considers subchannels assignment which are associated with discrete variables in the formulated problem, the problem is NP-hard. Moreover, even if the subchannel assignment information is known, because of the interference power resultant from the superposition coded signals of other users on a specific subchannel, the power allocation of the problem is non-convex [28]. As a result, joint subchannel assignment and power allocation of this problem can be considered as a mixed integer non-linear programming (MINLP) problem. Overall, joint optimization of both subchannel assignment and power allocation is not tractable for this case. Consequently, we solve this problem in two steps. Based on the assumption that the maximal power level of each user is subdivided equally among its allocated users, we apply many-to-many matching model [29], [30] to solve the subchannel-user mapping problem. Then, we apply iterative water-filling [27] and geometric programming (GP) [31] techniques to allocate power across the subchannels for different users. Iterative water-filling is a multi-user-channel power allocation technique, which is developed based on the insights of single-user water-filling solution. On the other hand, GP technique can solve special-form of non-convex problems using convex optimization solvers through variable transformation. Given the subchannel-user mapping information, our uplink power allocation problem is amenable to GP after applying some transformation on the objective function. Extensive simulation has been conducted to verify the effectiveness of our proposed uplink resource allocation scheme comparing with two very similar existing works [24], [26]. The results demonstrate that the proposed scheme always outperforms the existing works in terms of computational complexity, the usage of resource and overall performance.

The rest of the paper is organized as follows. Along with the background information and the description of the system, we formulate our uplink resource allocation problem in Section II. The detailed solution approach is provided in Section III. Followed by the simulation methodology, we evaluate the performance of our uplink resource allocation scheme in Section IV. Finally, Section V concludes the paper with some direction on future research.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this paper, we consider an uplink scenario of a cellular network, which has one base station. Time is divided into frames, and in each frame, the entire pre-assigned spectrum for the system is divided into N subchannels with equal bandwidth. The resultant subchannels are the elements of a set, denoted by N. There are M number of users in the system, and the corresponding set holding these users is denoted by M. Using the subchannels in set N as the transmission media, the users in set M transmit data to the base station. Each user m in set M has the maximal power level, denoted by $p_m^{\text{max}}$. Both the base station and the users in the system are equipped with NOMA technologies. The users transmit their data using superposition coding (SC) technique over a set of subchannels. Whereas, the receiver, i.e., the base station applies SIC technique on each subchannel to decode the
superimposed signals, and extracts the corresponding signal of each individual user. However, before the uplink transmission operation, it is required to schedule subchannels and power across the users optimally so that the capacity of the system is maximized. We assume that the scheduling scheme in the system is centralized, and the base station is appointed to conduct this operation. To develop this scheduling scheme, the entire channel state information (CSI) of the system is required, and hence the base station is aware of all these information. At the beginning of each time frame, all users send their CSI to the base station via some reliable control channels.

We consider the block fading channel model. It implies that the CSI of the subchannels in the system remains constant over a time frame, however vary independently across different time frames. Although NOMA techniques have various classification, we plan to exploit power-domain NOMA. We assume that the base station assigns \( M \) number of users on the \( n \) th subchannel, and the corresponding set holding these users is denoted by \( \mathbf{M}_n \). If each individual user \( m \) transmits \( \sqrt{p_m^n} s_m \) symbol on subchannel \( n \), the symbol received by the base station on this subchannel can be expressed as

\[
x_n = \sum_{m=1}^{M} \sqrt{p_m^n} s_m,
\]

where \( s_m \) is the modulated symbol of the \( m \)th user on subchannel \( n \), and \( p_m^n \) is the power level assigned to user \( m \) on subchannel \( n \). Consequently, the signal of user \( m \), received by the base station on subchannel \( n \), can be represented as

\[
y_m^n = h_m^n x_n + z_n = \sqrt{p_m^n} h_m^n s_m + \sum_{i=1,i \neq m} \sqrt{p_i^n} h_m^n s_i + z_n,
\]

where \( h_m^n \) is the channel gain of user \( m \) on the \( n \)th subchannel, \( z_n \) is the noise power over subchannel \( n \), which follows Additive White Gaussian Noise (AWGN) distribution with mean zero and variance \( \sigma^2 \), i.e., \( z_n \approx \mathcal{C}\mathcal{N}(0, \sigma^2) \). The noise power of subchannel \( n \) is statistically same for all users. In NOMA systems, each subchannel is shared by multiple users. Consequently, each user on subchannel \( n \) receives its signal as well as the interference signals from other users on the same subchannel. Therefore, without SIC at the base station, the received SINR of the \( m \)th user on subchannel \( n \) is given by

\[
\text{SINR}_m^n = \frac{p_m^n |h_m^n|^2}{\sigma^2 + \sum_{i=1,i \neq m} p_i^n |h_i^n|^2} = \frac{p_m^n g_i^n}{1 + \sum_{i=1,i \neq m} p_i^n g_i^n},
\]

where \( \sigma^2 = E[|z_n|^2] \) is the noise power on subchannel \( n \), and \( g_i^n = |h_i^n|^2/\sigma^2 \) is the normalized channel gain of user \( m \) on subchannel \( n \). Based on Shannon’s capacity formula, the sum-rate of subchannel \( n \) is given by

\[
R_n = \sum_{m=1}^{M} \log_2 (1 + \text{SINR}_m^n) = \sum_{m=1}^{M} \log_2 \left( 1 + \frac{p_m^n g_i^n}{1 + \sum_{i=1,i \neq m} p_i^n g_i^n} \right).
\]

In NOMA systems, the SIC process is implemented at the receiver to reduce the interference from other users on the same subchannel. According to [24], it has been proved that the optimal decoding order is equivalent to the decreasing order of received power. In this way, the interference imposed on each user by other users of the same subchannel is reduced, and consequently the sum-capacity and proportional fairness of the system are enhanced. Based on this truth, we adopt the following rule while determining the decoding order of the users transmitting on the same subchannel. As we mentioned in the introduction that our solution approach of this problem consists of two steps. In the first step, based on the assumption that the maximal power level of each user is equally subdivided among its allocated subchannels. In this case, we decode the users superimposed on a subchannel in the decreasing order of their received power.

For the second case, when the user and subchannel assignment information are known, the power allocation of the superimposed users should follow some order. From the insights of the decoding order idea of [24] as well as intuitively, it is obvious that the user with better gain in any subchannel should be assigned with larger transmit power. In this way, the interference imposed on any user (assigned to any subchannel) caused by other users is reduced. A sample example of this idea has been provided in Fig. 1.

![Fig. 1: A sample example for the organization of decoding order in subchannel n (b) is the organized version of (a)], where \( p_1^n g_1^n = 20, p_2^n g_2^n = 10 \) and \( p_3^n g_3^n = 30 \) (1st case), or \( g_1^n = 0.2, g_2^n = 0.1 \) and \( g_3^n = 0.3 \) (2nd case).
other all users assigned to a particular subchannel impose interference on a specific user. On subchannel \( n \), denote that the users in set \( M^m_n \) produce interference for user \( m \). Consequently, the sum-rate of subchannel \( n \) (in (4)) can be rewritten as

\[
R_n = \sum_{m=1}^{M^m} \log_2 \left( 1 + \frac{p^m_n g^m_n}{1 + \sum_{i \in M^m_n} p^i g^i} \right). \tag{5}
\]

In this work, while preserving the power constraints of all users, our objective is to allocate the subchannels in set \( N \) across all users in set \( M \), and assign power level to each subchannel-user slot so that the capacity of the system is maximized. Clearly, this is an optimization problem. To formulate this problem, we define a binary variable \( \alpha^m_n \). \( \alpha^m_n = 1 \) implies that subchannel \( n \) is assigned to user \( m \), and \( \alpha^m_n = 0 \) means the other case. It is proved in [26] that the more users are assigned to a subchannel, the better the system capacity is. However, due to the varying nature of wireless channels and the decoding complication of SIC technique, not necessarily more users assigned to a subchannel will enhance the system throughput. While giving weight to this observation and insight, we assume that maximum \( K \) number of users can be assigned to a subchannel. Therefore, the uplink subchannel and power allocation problem in this context can be formulated as follows.

\[
\max_{\{p^m_n, \alpha^m_n\}} \sum_{m \in M} \sum_{n \in N} \alpha^m_n \log \left( 1 + \frac{p^m_n g^m_n}{1 + \sum_{i \in M^m_n} p^i g^i} \right), \tag{6}
\]

subject to,

\[
\sum_{m \in M} \alpha^m_n \leq K, \quad \forall n \in N, \tag{7}
\]

\[
\sum_{n \in N} \alpha^m_n \leq \infty, \quad \forall m \in M, \tag{8}
\]

\[
\alpha^m_n \in \{0, 1\}, \quad \forall m \in M, \quad \forall n \in N, \tag{9}
\]

\[
\sum_{n \in N} \alpha^m_n p^m_n \leq p^\text{max}_m, \quad p^m_n \geq 0, \quad \forall m \in M. \tag{10}
\]

In the above formulation, there are two types of variables, i.e., \( \{\alpha^m_n\} \) and \( \{p^m_n\} \). \( \{\alpha^m_n\} \) are the set of discrete variables, and the problem is NP-hard because of these variables. On the other hand, even if the information of set \( \{\alpha^m_n\} \) is known, because of the interference term \( 1 + \sum_{i \in M^m} p^i g^i \) inside the log term of (6), the problem is non-convex. While considering the overall structure, we can say that the problem is jointly NP-hard and non-convex.

### III. Solution Approach

In this section, we explore the solution approach of the uplink resource allocation problem of a NOMA system described in the previous section. The entire problem is formulated in (6)-(10). Apparently, due to the discrete nature of subchannel assignment (i.e., variables \( \alpha^m_n \)) and the continuous nature of power assignment (i.e., variables \( p^m_n \)), this is a MINLP problem. This type of problem even in conventional OMA systems is intractable. When it comes to the case of NOMA systems, due to the superimposition of multiple users on the same subchannel, the solution of this problem brings further complication. Therefore, we have decomposed the problem in two parts. In the first part, based on the assumption that the maximal power level of a user is subdivided equally among its allocated subchannels, we solve the subchannel-user mapping problem. In this case, we find that two-sided matching model is appropriate to capture the structure of this problem. Since one user can be assigned with multiple subchannels and one subchannel can have multiple users, many-to-many matching scheme is expected to solve the first problem. For the solution of the second part of the problem, we assume that we have the subchannel-user mapping information. Even though the subchannel-user mapping information is known, the power allocation across the subchannels and users, given the power constraints of the users, is a non-convex problem. Based on the structure of the problem, we find that iterative water-filling and GP are very appropriate techniques to solve this problem.

#### A. Subchannel and User Mapping Scheme

In this system, intuitively, assignment of many users to a subchannel and allocating multiple users to a subchannel (to follow the guidelines of NOMA technique) is envisioned to enhance the overall throughput. Given the power constraint of each user, this problem is NP-hard. However, the nature of the problem implies that many-to-many matching model is appropriate to solve this problem. Given that maximum \( K \) number of users can be multiplexed on a subchannel, \( M \) users in set \( M \) and \( N \) subchannels in set \( N \) are two sets of players of this many-to-many matching relation. Note that each user \( m \) can have infinite \( (N \) in practice) number of users if possible. However, since user \( m \) has maximal power constraint \( p^\text{max}_m \), this should be subdivided equally among its allocated subchannels.

**Definition 1:** A many-to-many matching \( \mu \) is a mapping from set \( M \) to set \( N \) such that every \( m \in M \) and \( n \in N \) satisfy the following properties:

- \( \mu(m) \subseteq N \) and \( \mu(n) \subseteq M \)
- \( |\mu(m)| \leq \infty, \forall m \in M \)
- \( |\mu(n)| \leq K, \forall n \in N \)
- \( n \in \mu(m) \) if and only if \( m \in \mu(n) \)

where \( \mu(m) \) is the set of partners for user \( m \) and \( \mu(n) \) is the set of partners for subchannel \( n \) under the matching model \( \mu \). The definition states that each user in set \( M \) is matched to a subset of subchannels in set \( N \), and vice versa. In other words, each user may choose a set of subchannels as the communication media, whereas each subchannel may choose a set of users to be assigned with in order to maximize the overall benefit of the system. However, before accomplishing these assignment operations, each user needs to have preference list based on some criteria. The criterion of constructing preference list for users is based on their received power from the subchannels. For example, if the gain of subchannel \( n \) for user \( m \) is \( g^m_n \) and assigned power level of this subchannel is \( p^m_n \), then the received power from this subchannel is \( p^m_n g^m_n \). We use the notation \( \Omega^m_n \rightarrow \Omega^m_m \) to imply that user \( m \) wants to have the
subchannels in subset $\Omega'_m$ than the subchannels in subset $\Omega''_m$, where $\Omega'_m \subseteq \mathcal{N}$ and $\Omega''_m \subseteq \mathcal{N}$. Similar analogy can be made for any subchannel $n$ in set $\mathcal{N}$. Preference of each subchannel $n$ is based on the overall benefit (i.e., throughput) of the system. For example, if user $m$ chooses subchannel $n$, this subchannel only accepts this user if and only the system performance is enhanced by this allocation.

To solve our subchannel-user mapping problem, we are interested to look at a stable solution, in which there are no players that are not matched to one another but they all prefer to be partners. Since subchannel players give preference to the overall throughput of the system while choosing partners from set $\mathcal{M}$, stable solution is envisioned to be the optimal solution for this problem. In many-to-many matching models [30], many stability concepts can be considered depending on the number of players that can improve their utility by forming new partners among one another. However, due to the large number of players ($\mathcal{M} \cup \mathcal{N}$) in our problem, identifying optimal subset of partners for a player is intractable. Consequently, we choose to solve the matching problem by identifying partner one by one from the opposite set. This way of choosing partner in the matching model brings pair-wise stability.

In Definition 2 and Definition 3, we highlight some properties of pair-wise stable matching relation. For the sake of these definitions, we define some notations as follows. Faced with a set $\mathcal{N} \subset \mathcal{N}$ of possible partners, a player $m \in \mathcal{M}$ can determine which subset of set $\mathcal{N}$, it wishes to match to. We denote this choice set as $\mathcal{C}_m(\mathcal{N})$.

**Definition 2:** A matching relation $\mu$ is pairwise stable if there does not exist a pair $(m, n)$ with $m \notin \mu(n)$ and $n \notin \mu(m)$ such that $\phi \in \mathcal{C}_m(\mu(m) \cup \{n\})$ and $\varphi \in \mathcal{C}_n(\mu(n) \cup \{m\})$, and at the same time both $\{\phi\} \succ_m \mu(m)$ and $\{\varphi\} \succ_n \mu(n)$ are satisfied.

**Definition 3:** Let $\hat{\mathcal{M}}$ is the subset of users in set $\mathcal{M}$. The preference of subchannel $n$ is called substitutable if there exist users such that $m, m' \in \mathcal{C}_n(\hat{\mathcal{M}})$, then $m \in \mathcal{C}_n(\mathcal{M}/\{m'\})$ is satisfied.

While satisfying the properties of stable many-to-many matching relation, we have proposed an algorithm in Algorithm 2. Note that in this algorithm, we are interested in pair-wise stability, and hence the players (i.e., users and subchannels) choose their partners one-by-one instead of a subset. We have adopted a few paradigms or strategies in order to bring stability in this matching relation or enhance the overall system performance. The description of the algorithm is as follows. First, $\Omega_n, m \in \mathcal{M}$ are initialized with $\emptyset$, which basically contains the allocated subchannels of user $m \in \mathcal{M}$. At the same time, $\mathcal{M}_n, n \in \mathcal{N}$ are initialized with $\emptyset$ as well. Over the iterations, these sets are filled by the allocated subchannels and users, respectively. At the initialization phase, each user $m \in \mathcal{M}$ also constructs its subchannel preference list based on the descending order of their received power level. If the gain of subchannel $n$ for user $m$ is $g^m_n$ and the assigned power level is $p^m_n$, the received power level of this subchannel for this user is $p^m_n g^m_n$. Since we have an assumption that the maximal power level of user $m$ is subdivided equally among its allocated subchannels, the preference list of user $m$ is constructed based on the assumption that $\max_{\mathcal{M}_m} \Omega_m$ amount of power is reserved for subchannel $n \in \{\mathcal{N}/\Omega_m\}$. Then, inside the outer-most loop (between step 3 and step 34), if no assignment is possible inside the second outer loop (between line 4 and line 33), the algorithm terminates. Inside the inner-most loop (between step 5 and step 32), each user $m$ chooses its most preferred unallocated subchannel $n$. At this point, two conditions are possible. The first condition is that the number of allocated users on subchannel $n$ can be less than $K$ (maximum allowable number of users per subchannel), and the second condition is the other case. If the first condition is true, we can apply two strategies for this subchannel-user assignment: either user $m$ is substituted by one of the existing users (e.g., $n' \in \mathcal{M}_m$) on subchannel $n$, or user $m$ can be added to this subchannel. Each of these strategies is inserted to strategy set $\mathcal{S}$ (which was initialized before initiating the loop). Whereas, for the second case, only addition strategy is possible. After filling the strategy set $\mathcal{S}$ no matter the number of allocated users on subchannel $n$ is less than or equal to $K$, the elements of $\mathcal{S}$ are filtered based on some criterion, which is as follows. If strategy $s$ is a substitution policy, let $m'$ be to-be-replaced user, and hence $\mathcal{N}' = \Omega_m \cup \Omega_{m'}$ is the set of affected subchannels. Moreover, let $thrv$ be the total computed throughput (following (5)) of the subchannels in set $\mathcal{N}'$ before applying strategy $s$. Then, after applying strategy $s$ and adjusting the power level of user $m$ and $m'$ in set $\mathcal{N}'$, in the similar manner, the total throughput is computed (denoted by $thrv$). Strategy $s$ is only added to set $\mathcal{CS}$ if and only if this throughput (due to applying this strategy) is bigger than $thrv$. For the addition strategy, the affected subchannels are the ones in set $\Omega_m$, and hence $\mathcal{N}' = \Omega_m$. For this strategy, in the similar manner, set $\mathcal{S}$ is filtered and set $\mathcal{CS}$ is updated. Finally, $s_{Best}$ strategy is chosen based on the total throughput each strategy incurs. If $s_{Best}$ is empty, the inner-most loop continues, and the next user is chosen from set $\mathcal{M}$ for building its possible strategy set. If $s_{Best}$ is not empty, the corresponding strategy is executed. As a result, set $\Omega_m$, set $\Omega_{m'}$ (only for the substitution strategy), and set $\mathcal{M}_n$ are updated. The power level of user $m$ for the subchannels in set $\Omega_m$ are adjusted, and its preference list is updated as well due to the updated power level. For only substitution strategy, the power level of user $m'$ in its affected subchannels and its preference list is updated. By analyzing the algorithm, we conclude Proposition 1, Lemma 1 and Theorem 1.

**Proposition 1:** Rejected users by the subchannels are not final. For example, if a user $m \in \mathcal{M}$ is rejected by subchannel $n \in \mathcal{N}$ at some iteration $i$, at the $i'$th ($i' \geq i$) iteration, it is possible that both $m \in \mu(n)$ and $n \in \mu(m)$ will appear true.

1At this point, it is assumed that the system has reached a stable situation or the performance improvement is no longer possible.
Algorithm 1  The uplink subchannel-user mapping algorithm using many-to-many matching model.

1: $\Omega_m \leftarrow \emptyset, \forall m \in M; M_n \leftarrow \emptyset, n \in N.$
2: Each user $m \in M$ produces its preference list.
3: repeat
4: for $i \leftarrow 1$ to $N$ do
5: for $m \in M$ do
6: $n \leftarrow$ The best unallocated subchannel from the preference list of user $m$.
7: $S \leftarrow \emptyset$, $CS \leftarrow \emptyset$, and $\text{Thruput} \leftarrow \emptyset$.
8: if The number of assigned users on subchannel $n$ is less than $K$ then
9: Construct each strategy $s$ that is supposed to replace each user $m' \in M_n$ by user $m$, and insert the corresponding $s$ to set $S$.
10: Construct another strategy $s$ that is supposed to add user $m$ to set $M_n$, and insert the corresponding $s$ to set $S$.
11: else if The number of assigned users on subchannel $n$ is equal to $K$ then
12: Construct each strategy $s$ that is supposed to replace each user $m' \in M_n$ by user $m$, and insert the corresponding $s$ to set $S$.
13: end if
14: for $s \in S$ do
15: Determine the affected subchannel list $N'$ for strategy $s$.
16: Adjust the power level of user $m' \in M_n$, $\forall m' \in N'$.
17: $\text{thr} \leftarrow$ The sum-throughput of all subchannels in set $N'$ due to strategy $s$.
18: Insert strategy $s$ to $CS$ and insert $\text{thr}$ to set $\text{Thruput}$ if $\text{thr}$ is larger than the sum-throughput of affected subchannel list $N'$ before applying strategy $s$.
19: end for
20: $s_{\text{Best}} \leftarrow \arg\max_{s \in CS} \text{Thruput}(s)$.
21: if $s_{\text{Best}}$ is the replacement strategy then
22: $\Omega_m \leftarrow \Omega_m \cup \{n\}$, $\Omega_{m'} \leftarrow \Omega_{m'} \setminus \{n\}$, and $M_n \leftarrow (M_n / \{m'\}) \cup \{m\}$ if $m'$ is the to-be-replaced user and $m$ is the new user on subchannel $n$.
23: Adjust the power level of the subchannels in $\Omega_m$ and $\Omega_{m'}$, and update the preference lists of user $m$ and user $m'$.
24: else if $s_{\text{Best}}$ is the addition strategy then
25: $m$ is the new user on subchannel $n$.
26: $\Omega_m \leftarrow \Omega_m \cup \{n\}$, and $M_n \leftarrow M_n \cup \{m\}$ if $m$ is the new user on subchannel $n$.
27: Adjust the power level of the subchannels in $\Omega_m$, and update the preference list of user $m$.
28: end if
29: if $s_{\text{Best}}$ is not empty then
30: Terminate this loop.
31: end if
32: end for
33: end for
34: until The performance enhancement of the system is not possible.

Proof: Let assume $K = 2$, which implies that maximum allowable number of users per subchannel is 2. Consider that user 1 and user 2 are already matched with subchannel $n$ in some iteration $i$, and $p_2 > p_1$ holds. At iteration $i'$ ($i' > i$), user 3 has to come to obtain subchannel $n$ with the power level $p_3$, and $p_2 > p_3 > p_1$ holds. For the sake of simplicity, we further assume that $g_2^i = g_3^i = g_4^i = 1$. We know that if $p_2 > p_3 > p_1$ holds, $\log_2(1 + p_2^i + p_3^i) > \log_2(1 + p_2^i + p_1^i)$ always satisfies. Therefore, at this stage, based on the replacement strategy our algorithm provides, user 3 is replaced by user 1. Since user 1 is unallocated from subchannel $n$, in other subchannels (of set $\Omega_1$) to which user 1 is belong to, the power level of this user will be increased. This is because the algorithm ensures that the maximal power level $p_1^{\text{max}}$ is equally subdivided among the allocated subchannels of user 1. In the similar manner and for the same reason, user 1 may further be replaced by some other user on its some other assigned subchannel (in set $\Omega_1$). At this stage, at iteration $i''$ ($i'' > i'$), user 1 is able to compete (although rejected already at iteration $i'$) for obtaining the $n$th subchannel again with the increased level of power, denoted by $p_1^{i''}(p_1^{i''} > p_1^i)$. Thus, we prove that user 1 may come again to choose subchannel $n$, and can be replaced by user 3. This is due to the fact $\log_2(1 + p_2^i + p_3^i) > \log_2(1 + p_2^{i''} + p_3^{i''})$ because of $p_1^{i''} > p_3^i$. By adopting the strategy in this proposition, the algorithm ensures as better performance as possible for the system.

Lemma 1: The subchannel-user mapping algorithm (i.e., Algorithm 1) is guaranteed to converge to a pair-wise stable matching relation.

Proof: We prove this lemma by contradiction. Suppose that there exist a user $m$ and a subchannel $n$ with $m \not\in \mu(n)$ and $n \not\in \mu(m)$ such that $\phi \in C_{\mu(n)}(\mu(m) \cup \{n\}), \phi \in C_{\mu(n)}(\mu(m) \cup \{n\})$, and at the same time both $\phi \not\geq_{\mu(n)}$ and $\phi \not\geq_{\mu(m)}$ are satisfied. Since $\phi \not\geq_{\mu(n)}$ and $\phi \not\geq_{\mu(m)}$ is true, user $m$ must propose subchannel $n$ in some earlier iteration to be paired with. However, at the same time, both $m \not\in \mu(n)$ and $n \not\in \mu(m)$ are true. Consequently, at the proposal time of user $m$, either subchannel $n$ had some better preference compared with user $m$ and rejected this user, or accepted this user and then made a replacement with some other user in the latter iteration. Therefore, $m \not\in C_{\mu(n)}(\mu(m) \cup \{n\})$ cannot be a false statement, and hence matching relation $\mu$ cannot be unstable.

Theorem 1: Algorithm 1 terminates after finite number of iterations.

Proof: Algorithm 1 is proposed to solve the subchannel-user mapping problem, and this is an optimization problem. No matter the problem is convex or non-convex, we know that every optimization problem has a unique global solution. If the problem is non-convex, there might be some local optimal solutions. However, for the convex problem, the solution is unique, which can be assumed as both local and global. As we verify before that the problem discussed herein is non-convex and NP-hard, the global optimal solution requires to search all possible feasible solution spaces, which is computationally intensive and is not feasible to implement in a wireless system, where channel turnaround time is in millisecond/microsecond level. Consequently, we proposed Algorithm 1 to solve this problem. We do not claim that the algorithm always finds...
the global optimal solution as this is based on the many-to-
many matching scheme. In the many-to-many matching model,
since the number of players (i.e., users and subchannels) on
the both sides is large, in terms of their associations, many
combinations are possible as mentioned above. Therefore, we
mostly focused on pair-wise stability, and Algorithm 7 is de-
veloped based on this concept. The outer-most loop continues if
and only if at least one allocation (subchannel-user mapping)
is executed. Interestingly, in our algorithm, every allocation
enhances the system throughput compared to the throughput
before that particular allocation. Step 18 of the algorithm is the
evidence of this statement. Through addition and substitution
operations, for a user, we define possible strategies for the
tagged user and its preferred subchannel. This implies that the
tagged user will be added to its most preferred subchannel or
will replace an existing user of that subchannel if only if the
system capacity is enhanced by this allocation. Therefore, in
every allocation, if the system capacity is enhanced, eventually,
the process approaches the convergence state as the solution
of the problem exists in the finite domain. Moreover, no matter
the problem is convex or non-convex, the global or local optimum point has convex nature. Consequently, after finite
number of iterations, the algorithm converges and terminates.
In order to provide practical evidence of the convergence
event, we have plot Fig. 2. In each subfigure of this figure,
we show the system throughput with the increasing outer-
most loop iterations for different values of $K$. As observed,
in each iteration, the system throughput is increased little by
little before reaching the convergence state.

In Fig. 3a and Fig. 3b we present the average number of
iterations the outer-most loop runs before reaching the optimal
point with the increasing number of users and the increasing
value of $K$, respectively. The total number of subchannels
in both figures is constant and fixed. Intuitively, given the
number of users and the number of subchannels to be matched,
the outer-most loop should run the times surrounding the
value of $K$. However, when the number of users is less, not
necessarily, each subchannel obtains exactly $K$ number
of users. Moreover, when the number of users in the system
is less, the number of strategies with substitution operation
is relatively less compared to the case when the number of
users in the system is higher. In this case, the outer-
most loop terminates in less number of iterations, which is
obvious in Fig. 3a. On the other hand, when the value of $K$
is lower, for the given number of users and subchannels in
the system, each subchannel obtains relatively less number of
users compared to the case with larger value of $K$. Therefore,
due to the restriction on the less number of users allocated to
each subchannel, the number of strategies accompanied with
substitution and addition operations is less in this case as
well similar to the other case. Consequently, the outer-most
loop requires less number of iterations to run before reaching
the optimal possible point of the system. Fig. 3b presents the
corresponding observation in this context.

**Computational Complexity of Algorithm 7** The joint worst
case complexity of the inner two loops of the algorithm is $O(MN)$. Inside these loops, all operations occur in constant
time, and so we can ignore the complexity of these operations.

**Algorithm 1**

\[
\begin{align*}
\text{1)} & \quad \text{Initialize } (m, n) \in \mathbb{M}^n, \quad \text{subject to } m \leq K, \quad n \leq K \\
\text{2)} & \quad \text{if at least one allocation is executed, then continue, else terminate} \\
\text{3)} & \quad \text{for } (m, n) \in \mathbb{M}^n \\
\text{4)} & \quad \text{repeat} \\
\text{5)} & \quad \text{for } m \leq K \text{ and } n \leq K \\
\text{6)} & \quad \text{for } (m, n) \in \mathbb{M}^n \\
\text{7)} & \quad \text{if } \text{the system throughput is enhanced, then continue, else terminate} \\
\text{8)} & \quad \text{end for} \\
\text{9)} & \quad \text{end for} \\
\text{10)} & \quad \text{end for} \\
\text{11)} & \quad \text{end repeat} \\
\text{12)} & \quad \text{end for} \\
\end{align*}
\]

Fig. 2: Increasing system throughput as the outer-most loop iteration of Algorithm 7 advances.

Fig. 3: The number of outer-most loop iterations required before achieving stability.

Mainly, the running time of the outer-most loop dominates
the computation time of the entire algorithm. As shown in
Fig. 3a and Fig. 3b, the iterations of this outer-most loop is
proportional to the number of users ($M$) and the maximum
allowable number of user per subchannel ($K$), but this is not
a very large number. Therefore, we conclude that the algorithm
has polynomial time complexity. Whereas, the complexity
of the brute-force searching operation to obtain the optimal
solution is $O(M^N \times (M-1)^n \times \cdots) \approx O(M^N)$, which is an order
of exponential series.

**B. Power Allocation Schemes**

From Algorithm 7 we know the subchannel-user mapping
information, i.e., $\mathbb{M}_n, n \in \mathbb{N}$ and $\Omega_m, m \in \mathbb{M}$. This informa-
tion is derived based on the assumption that the maximal
power level $p_{\text{max}}^n$ of user $m$ is equally subdivided among
its allocated subchannels, i.e., $p_{\text{max}}^n = \frac{p_{\text{max}}^m}{|\Omega_m|}, m \in \mathbb{M}$. However, in (6), we see that the instantaneous rate of user $m$
on subchannel $n$ is positively proportional to $p_{\text{max}}^n$ and inversely
proportion to the interference power level of other users, i.e.,
$p_{i}^n, i \in \mathbb{M}_{\text{m}}$. Consequently, even if the information about
$\mathbb{M}_n, n \in \mathbb{N}$ and $\Omega_m, m \in \mathbb{M}$ is known, the power allocation
across all subchannel-user slots, i.e., $p_{\text{max}}^n, m \in \mathbb{M}, n \in \mathbb{N}$, is
an optimization problem. Consequently, the next objective of
this resource allocation scheme is to allocate power across
all subchannel-user slots optimally. We have adopted two
approaches to solve this power allocation problem, the individual
description of which is provided in the following discussions.

1) Iterative Water-Filling Algorithm: Given that
subchannel-user mapping information, $\mathbb{M}_n, n \in \mathbb{N}$ and
$\Omega_m, m \in \mathbb{M}$ are known, the power allocation problem can be
written as

\[
\begin{align*}
\text{1)} & \quad \text{Initialize } (m, n) \in \mathbb{M}^n, \quad \text{subject to } m \leq K, \quad n \leq K \\
\text{2)} & \quad \text{if at least one allocation is executed, then continue, else terminate} \\
\text{3)} & \quad \text{for } (m, n) \in \mathbb{M}^n \\
\text{4)} & \quad \text{repeat} \\
\text{5)} & \quad \text{for } m \leq K \text{ and } n \leq K \\
\text{6)} & \quad \text{for } (m, n) \in \mathbb{M}^n \\
\text{7)} & \quad \text{if } \text{the system throughput is enhanced, then continue, else terminate} \\
\text{8)} & \quad \text{end for} \\
\text{9)} & \quad \text{end for} \\
\text{10)} & \quad \text{end for} \\
\text{11)} & \quad \text{end repeat} \\
\text{12)} & \quad \text{end for} \\
\end{align*}
\]
\[
\max_{\{p^n_m\}_{m \in M, n \in N}} \sum_{n \in N} \log_2 \left(1 + \sum_{m \in M} p^n_m g^n_m\right),
\]
\[\text{s.t., } \sum_{n \in \Omega_m} p^n_m \leq p^\text{max}_m, \forall m \in M, \tag{11}\]

where the objective function is the simplified version of the objective function in (6). This is actually a multi-user water-filling problem. Using the dual decomposition method [28], the solution of this problem is described as follows. Taking the dual variables \(\lambda_m, m \in M\), the Lagrangian of the problem in (11) can be written as

\[
L((p^n_m)_{m \in M, n \in N}) = \sum_{n \in N} \log_2 \left(1 + \sum_{m \in M} p^n_m g^n_m\right),
\]
\[+ \sum_{m \in M} \lambda_m \left(p^\text{max}_m - \sum_{n \in \Omega_m} p^n_m\right). \tag{12}\]

Taking the derivative of the Lagrangian in (12) with respect to \(p^n_m, m \in M, n \in N\), we obtain

\[
p^n_m = \frac{1}{\lambda_m} - \frac{1}{g^n_m} \left(1 + \sum_{i \in M, i \neq m} p^n_i g^n_i\right), \forall m \in M, \forall n \in N. \tag{13}\]

If we compare the solution in (13) with the single-user water-filling solution \(p^n_m = 1/\lambda_m - 1/g^n_m\), it is obvious that the optimal power level of one user considers the received power of other users as noise. Based on this intuition, we develop an iterative algorithm in Algorithm 2 to solve this power allocation problem. The algorithm works as follows. First, power level of all users over all subchannels are initialized. Then, for each user, water-filling power allocation problem is solved assuming the power level of other users as noise. Once the single-user water-filling solutions are obtained for all users, the resultant solutions of all users are replaced by the previously initialized power levels. This operation is continued until the performance of the system appears to be saturated.

While solving the single-user water-filling problem in Algorithm 2 typically, bisection search is applied to obtain the optimal value of \(\lambda_m, m \in M\). If the proper interval of the bisection search is not chosen, running time of the bisection search is huge. Moreover, the accuracy of the solution obtained from the bisection search is greatly dependent on the precision level of \(\lambda_m\) as this is a variable with continuous nature. Therefore, to obtain the optimal value of \(\lambda_m\), we have developed a low complexity procedure in Algorithm 3. In this algorithm, for user \(m\), \(A_m\) is a vector, the elements of which are \(\frac{1}{g^n_m} \left(1 + \sum_{i \in M, i \neq m} p^n_i g^n_i\right), n \in \Omega_m\). The insights of this procedure is developed based on the following relation for individual user \(m\)

\[
\sum_{n \in \Omega_m} \frac{1}{\lambda_m} - \frac{1}{g^n_m} \left(1 + \sum_{i \in M, i \neq m} p^n_i g^n_i\right) = p^\text{max}_m \tag{14}\]
\[\lambda_m = \frac{p^\text{max}_m + \frac{1}{g^n_m} \left(1 + \sum_{i \in M, i \neq m} p^n_i g^n_i\right)}{D_m}. \tag{15}\]

**Algorithm 2 The iterative water-filling algorithm to calculate optimal \(\{p^n_m\}_{m \in M, n \in N}\).**

1: \(p^n_m \leftarrow 0, \forall m \in M, \forall n \in N\).
2: **repeat**
3:   **for** \(m \in M\) **do**
4:     \(p^n_m \leftarrow \arg \max_{p^n_m \in \Omega_m} \sum_{n \in \Omega_m} \log_2 \left(1 + \sum_{i \in M, i \neq m} p^n_i g^n_i\right)\),
5:     **s.t.,** \(\sum_{n \in \Omega_m} p^n_m \leq p^\text{max}_m\).
6:   **end for**
7: **until** The performance improvement is not possible.

**Algorithm 3 The iterative process to calculate optimal \(\lambda_m\) for user \(m\) in Algorithm 2.**

1: The elements of \(A_m\) is sorted in the ascending order.
2: \(j \leftarrow 1, G_n \leftarrow 0, G_d \leftarrow 0\).
3: **repeat**
4:   \(G_n \leftarrow G_n + 1\).
5:   \(G_d \leftarrow G_d + A_m(j)\).
6:   \(\lambda_m(j) \leftarrow \frac{G_n}{G_d}\).
7:   **if** \(\lambda_m(j) \geq 1/A_m(j + 1)\) **then**
8:     \(\lambda^*_m \leftarrow \lambda_m(j)\).
9:   **break** the loop.
10: **end if**
11: \(j \leftarrow j + 1\).
12: **until** \(j \leq |\Omega_m|\).

2) **Geometric Programming:** Another way to write the problem in (11) is as follows

\[
\min_{\{p^n_m\}_{m \in M, n \in N}} \frac{1}{\prod_{n \in N} \left(1 + \sum_{m \in M} p^n_m g^n_m\right)} \approx \min_{\{p^n_m\}_{m \in M, n \in N}} \frac{1}{\prod_{n \in N} \left(1 + \sum_{m \in M} p^n_m g^n_m\right)},
\]
\[\text{s.t., } \sum_{n \in \Omega_m} p^n_m \leq p^\text{max}_m, \forall m \in M. \tag{16}\]

GP is an optimization technique that can solve some non-convex problem by adopting some transformation on the optimization variables. The objective and constraint functions with which GP deals are posynomials and monomials. The objective function in (16) is the ratio of two posyn-
omials. The ratio of two posynomials is not a posynomial, and hence this problem is still not amenable to GP. However, there are some heuristics, such as single condensation method, double condensation method [31] that can be used to make the problem amenable to GP. We have adopted single condensation method to solve this problem. According to this method, the denominator (which is a posynomial) has to be approximated by some monomial. If we denote the denominator of the optimization problem by \( G(\{p_m^n\}_{m \in M_n, n \in N}) \), the approximated monomial of this function is

\[
G(\{p_m^n\}_{m \in M_n, n \in N}) = \prod_{n \in N} \left(1 + \sum_{m \in M_n} p_m^n a_m^n\right)
\]

\[
\approx \lambda \prod_{m \in M_n, n \in N} (p_m^n a_m^n), \quad (17)
\]

where \( a_m^n, m \in M_n, n \in N \) and \( \lambda \) are auxiliary variables. Given the values of \( \{p_m^n\}_{m \in M_n, n \in N} \), the values of the auxiliary variables can be obtained as follows.

\[
a_m^n = \frac{p_m^n}{G(\{p_m^n\}_{m \in M_n, n \in N})} \frac{\partial G(\{p_m^n\}_{m \in M_n, n \in N})}{\partial p_m^n}, \quad (18)
\]

\[
\lambda = \frac{G(\{p_m^n\}_{m \in M_n, n \in N})}{\prod_{m \in M_n, n \in N} (p_m^n a_m^n)}, \quad (19)
\]

and

\[
\frac{\partial G(\{p_m^n\}_{m \in M_n, n \in N})}{\partial p_m^n} = g_m^n \prod_{n' \in N, n' \neq n} (1 + \sum_{m' \in M_{n'}} p_{m'}^n g_{m'}^n). \quad (20)
\]

Given some initial values of \( \{p_m^n\}_{m \in M_n, n \in N} \), we require an iterative process in order to obtain the optimal values of these variables step by step. The steps of this iterative process are provided in Section III of [26]. The final values of the variables \( p_m^n, m \in M_n, n \in N \), obtained in the last iteration of the iterative process, is the solution of our defined optimization problem.

**IV. Performance Evaluation**

In this section, we evaluate the performance of our proposed uplink resource allocation scheme via extensive simulation. The detailed system setup and simulation settings are provided in the following subsection. Then, we exhibit the results obtained from the conducted simulation to verify the effectiveness of our scheme.

A. Simulation Setup

The cellular network, that we consider, has a circle-like shape. Since our proposed resource allocation scheme did not capture the interference from neighboring cells, we assume that the cellular network is isolated. The base station is located at the center of the cell, and the users are uniformly distributed in a circular range with 500 m radius. We set the minimum distance between users to 40 m, and the minimum distance from the users to the base station to 50 m. As mentioned previously, time is divided into frames. Each time frame is equivalent to 1 s, and during this frame, spectrum is subdivided among 20 subchannels and these are available to be allocated among \( M \) users.

Each subchannel is assumed to have 200 KHz bandwidth. According to [37], the theoretical limit of the channel capacity is given by \( \frac{-1}{ln(5P_b)} \), where \( P_b \) denotes the Bit Error Rate (BER). BER for the channel is configured as \( 10^6 \). The channel between the base station and a user is affected by shadow and Rayleigh fading. Shadowing effect follows log-normal distribution with variance 3.76. In order to calculate lognormal shadowing effect of a subchannel, we assume the reference distance as 1 km and the SNR for this reference distance is 28 dB. Reference shadowing effect has also the log normal distribution with variance 3.76. Rayleigh fading effect for each user over a subchannel follows Rayleigh distribution with zero mean and 10 scale factor. Using all these parameters, the gain of each subchannel for a user towards the base station is computed following (22) in [24]. The maximal power level of all users is set to 30 W. SC-coded signal on each subchannel at the base station is decoded following the SIC technique in [12, 88].

In addition to implement our proposed resource allocation scheme, we have implemented relevant other algorithms [24, 20] proposed in the literature already. For example, in [24], the authors proposed two heuristics in order to maximize the overall capacity and proportional fairness across the system. In the figures demonstrated in the following subsection, these are referred by Alg. 1 [24] and Alg. 2 [24]. The algorithm proposed in [20] is referred by Alg. [20]. Since we have adopted two techniques in order to allocate power to all subchannels across all users, while referring our algorithm, we use IWF and GP for iterative water-filling and GP power allocation schemes, respectively. Furthermore, obtaining the optimal subchannel-user mapping information is computationally intensive for a large-scale system, and hence we apply brute-force search for a system with 10 and 20 users. In the following subsection, for each data point, we conduct the simulation over 10000 time frames.

B. Simulation Results

In Fig. [4] we show the increasing system capacity with the increasing number of users. This is a natural trend. The more the number of users in the system the more the overall capacity. In this figure, we set \( K \) to 6. No matter the number of users in the system, our proposed scheme always outperforms existing other algorithms. Actually the algorithms proposed in [24] have an assumption that each user can get only one subchannel, which is the main reason of such degraded performance compared to other algorithms. If one user obtains only one subchannel, that assumption fails to exploit the multiuser-channel diversity of wireless systems, and consequently the overall capacity of the system is much lower. On the other hand, the total capacity obtained by the algorithm that is designed to maximize the overall capacity should be larger.

\footnote{The ratio of a posynomial and a monomial is a posynomial.}
Consider the power level of other users as interference level the capacity of that subchannel. This is because SIC technique relatively better level of power, it does not necessarily enhance. However, if a subchannel has a number of allocated users with user implies, that user obtains more power on that subchannel. Better gain of a subchannel for a nature and design, it is obvious that the algorithm only keeps compared to other subchannels. On the other hand, which has relatively better gain can have larger number of subchannels. Then, each user only wants to obtain its most preferred subchannel, and the corresponding subchannel either adds this user or substitutes the existing user if and only if the sum-capacity of that subchannel and other affected sub-channels is enhanced. Note that while solving the subchannel-user mapping problem, it is assumed that each user subdivides its maximal power equally among its allocated subchannels. Furthermore, even if the value of $K$ is large, exactly $K$ number of users allocated to a subchannel (especially the one with worse channel) not necessarily enhances the system throughput. The algorithm adds additional user to a subchannel if that user improves the sum-throughput of that subchannel and other affected sub-channels. All these design mechanisms allow the algorithm to overcome the drawbacks of the algorithm in [26], and outperforms it. Although for the power allocation of our scheme, we have adopted two techniques, iterative water-filling algorithm outperforms GP technique. Moreover, the way we have implemented iterative water-filling algorithm such that it has much less computational complexity compared to the other one. Therefore, we recommend iterative water-filling as the power allocation technique for our proposed resource allocation scheme.

For $M = 40$ in the system, Fig. 5 depicts the increasing system capacity with the increasing value of $K$. It is apparent that the more the number of users in a subchannel, the more enhanced the system capacity is. However, not necessarily the same $K$ number of users on every subchannel enhances the system capacity. This phenomenon particularly happens for the subchannel which has worse condition due to the subdivision of limited power level of each user among its allocated subchannels. Therefore, our algorithm assigns less number of users to some subchannels whenever necessary. However, the algorithm in [26] blindly assigns exactly $K$ number of users to each subchannel without giving attention to the performance of the system. This statement has been well-proved in this figure.

The evidence that our algorithm not necessarily assigns exactly $K$ number of users to all subchannels is strengthened in Fig. 6. In this figure, $K$ is set to 8. The subchannel which has relatively better gain can have larger number of users compared to other subchannels. On the other hand, the algorithm in [26] always assigns $K$ number of users to all subchannels. The algorithms in [24] assign the user to a subchannel which has the lowest interference level and while decoding the signal of one user. As a result, such mechanism of the algorithm does not enhance the system capacity. There must be some other algorithm that should solve the subchannel-user mapping problem in such a way that the drawback of this algorithm can be overcome.

Consequently, we have proposed an efficient algorithm to solve the subchannel-user mapping problem based on the many-to-many matching model. The algorithm is designed in such a way that in every iteration, the capacity is enhanced little by little. The algorithm terminates only when the performance of the system cannot be enhanced anymore. Based on the first requirement of many-to-many matching model, each user constructs its preference list in the descending order of the received power achieved from the subchannels. Then, each user only wants to obtain its most preferred subchannel, and the corresponding subchannel either adds this user or substitutes the existing user if and only if the sum-capacity of that subchannel and other affected sub-channels is enhanced.

![Graph showing comparison of system throughput with increasing number of users](image1)

**Fig. 4:** Comparison of system throughput with the increasing number of users.

![Graph showing comparison of system throughput with increasing maximum allowable number of users on a subchannel](image2)

**Fig. 5:** Comparison of system throughput with the increasing value of $K$.

than that obtained by the algorithm which is designed to maximize the proportional fairness. Fairness of the system is always achieved by sacrificing the capacity of the system.

The algorithm proposed in [26] has very close performance compared to ours. This algorithm is designed based on iterative water-filling algorithm. The idea of the algorithm is as follows. First, it is assumed that all users are allocated to all subchannels. On this setup, the iterative water-filling algorithm is applied for the power allocation. Obviously, the subchannel-user slot which has the worst gain, obtains the least amount of power. Consequently, the corresponding user is unallocated from the corresponding subchannel. Then, again iterative water-filling algorithm is applied until the constraint that, each subchannel obtains exactly $K$ number of users, is met. At this point, this algorithm is terminated. From the nature and design, it is obvious that the algorithm only keeps the users to a subchannel which have relatively better gain compared to other users. Better gain of a subchannel for a user implies, that user obtains more power on that subchannel. However, if a subchannel has a number of allocated users with relatively better level of power, it does not necessarily enhance the capacity of that subchannel. This is because SIC technique considers the power level of other users as interference level.

...
each user only obtains one subchannel. Therefore, these two algorithms also assign less number of users compared to the value of $K$.

In Fig. 7 and Fig. 8, we provide some detailed information about our algorithm to justify its superiority. In these figures, $K$ is set to 8 as well. Since our algorithm is based on the consideration to maximize the system capacity, the user who has the best gain on all subchannels, should have the highest throughput. At the same time, since Rayleigh fading effect is statistically similar for all users over all subchannels, the user closest to the base station should have the highest throughput, and so thus observed in Fig. 7. There are some exceptions as well due to the random nature of Rayleigh fading effect, such as users 9, 10, 11 and 12. Other than some exceptions, the throughput of the users have the decreasing trend with the increasing distance from the base station. Due to the aforementioned justifications, other algorithms incur less throughput for all users except the one in [26]. Previously, we claimed that our proposed algorithm is not necessarily globally optimal, and therefore the algorithm in [26] incurs larger throughput for one or two users. However, it is obvious in Fig. 7 that the throughput of more number of users incurred by our algorithm are better compared to the one in [26]. One of
the design insights of our proposed resource allocation scheme is to exploit multi-user-channel diversity of wireless systems. Consequently, by assigning more number of subchannels to each user, it is possible to enhance the system throughput, and so thus our algorithm does (as depicted in Fig. 8). However, the algorithms in [24] assign only one subchannel to a user, and hence their performance is much worse compared to others.

In Fig. 9 and Fig. 10 we further justify that our algorithm is superior in terms of resource efficiency as well. In these two figures, we plot the ratio of total subchannel-user slots and the number of used subchannel-user slots with the increasing number of users and with the increasing value of $K$, respectively. If the number of used subchannel-user slots is denoted by $D$, then the metric on the $Y$-axis of these two figures is $\frac{\sum_{n \in N} |M_n|}{D}$. It is obvious that the more the number of users and the larger the value of $K$, the more subchannels are used to support more users and to enhance the system throughput. Previously, we observed that the overall throughput achieved by the algorithm in [20] has close performance to ours, however using more resource slots in the system as depicted in Fig. 9 and Fig. 10. On the other hand, although the algorithms in [24] have the highest resource efficiency, they have the worst overall performance as proved in the previous results. These algorithms incur the highest performance in this case because of assigning only one subchannel to each user. Whereas, to exploit multi-user-subchannel diversity of wireless systems, our scheme assigns more subchannels to the users, and consequently achieves the highest overall performance.

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed an uplink resource allocation scheme of a NOMA system, in which the spectrum is divided into multiple subchannels. The objective of our resource allocation scheme is to allocate power and subchannels across the users of the system. Due to the discrete nature of subchannels and the characteristics of NOMA systems, the problem is NP-hard and non-convex. Since the optimal solution of the problem is intractable, we solved the problem in two steps. First, the subchannel-user mapping problem was solved using many-to-many matching model. Then, iterative water-filling algorithm and GP technique were applied to allocate power optimally across all assigned subchannels and users of the system. We conducted extensive simulation to verify the effectiveness of our proposed resource allocation scheme comparing with other existing works in the literature.

One of our previous experience [39] while dealing with interference is, reducing the allocated power level on the subchannels may bring better performance for the system compared to the case when the maximal power of the users are used. This is what exactly observed while allocating subchannels and power among the users in an OFDMA-based network surrounded by many neighboring cells. In such systems, transmission of users in one cell causes interference for the transmission of users in other cells. Since in NOMA systems, multiple users are superimposed on the same sub-channel, power level of one user may cause interference for other users. Both water-filling algorithm and GP technique use full power of a user to its allocated subchannels. However, we believe that using less power, it might be possible to enhance the system performance in terms of both capacity and resource efficiency. Another assumption of our work is, the network for which we proposed the uplink resource allocation scheme, is isolated and does not have any neighboring cell. Such assumption of the network is equivalent to ignoring the interference from neighboring cells. However, interference from the neighboring cells is a crucial factor especially for the uplink case in this context. As of our future work, we would like to continue the research in this direction in order to obtain more promising and useful results.

REFERENCES

[1] S. Tekinay, “Next generation wireless networks,” Springer US, 2002.
[2] S. C. Borst, A. Buaneswari, L. M. Drabeck, M. J. Flanagan, J. M. Graybeal, G. K. Hampel, M. Haner, W. M. MacDonald, P. A. Polakos, G. Rittenhouse, I. Saniee, A. Weiss, and P. A. Whiting, “Dynamic optimization in future cellular networks,” Bell Labs Technical Journal, vol. 10, pp. 99–119, 2005.
[3] A. Agarwal and K. Agarwal, “The next generation mobile wireless cellular networks: 4G and beyond,” American Journal of Electrical and Electronic Engineering, vol. 2, no. 3, pp. 92–97, 2014.
[4] A. Ali, W. Hamouda, and M. Uysal, “Next generation M2M cellular networks: Challenges and practical considerations,” CoRR, vol. abs/1506.06216, 2015. [Online]. Available: http://arxiv.org/abs/1506.06216
[5] S. Ghosh, K. Basu, and S. K. Das, “An architecture for next-generation radio access networks,” IEEE Netw., vol. 19, no. 5, pp. 35–42, Sept 2005.
[6] V. Suryaprakash and I. Malanchini, “Reliability in future radio access networks: From linguistic to quantitative definitions,” in Proc. IEEE IWQoS, June 2016, pp. 1–2.
[7] T. Edler and S. Lundberg, “Energy efficiency enhancements in radio access networks,” Ericsson Review, vol. 21, no. 1, 2014.
[8] K. Higuchi and Y. Kishiyama, “Non-orthogonal access with random beamforming and intra-beam SIC for cellular MIMO downlink,” in Proc. IEEE VTC Fall, 2013, pp. 1–5.
[9] Y. Endo, Y. Kishiyama, and K. Higuchi, “Uplink non-orthogonal access with MMSE-SIC in the presence of inter-cell interference,” in Proc. ISWCS, Aug 2012, pp. 261–265.
[10] J. Umehara, Y. Kishiyama, and K. Higuchi, “Enhancing user fairness in non-orthogonal access with successive interference cancellation for cellular downlink,” in Proc. IEEE ICCS, Nov 2012, pp. 324–328.
[11] N. Otao, Y. Kishiyama, and K. Higuchi, “Performance of non-orthogonal access with SIC in cellular downlink using proportional fair-based resource allocation,” in Proc. ISWCS, Aug 2012, pp. 476–480.
[12] N. I. Miridakis and D. D. Vergados, “A survey on the successive interference cancellation performance for single-antenna and multiple-antenna OFDM systems,” IEEE Commun. Surveys Tutorials, vol. 15, no. 1, pp. 312–335, Jan 2013.
[13] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, “System-level performance evaluation of downlink non-orthogonal multiple access (NOMA),” in Proc. IEEE PMRC, June 2013, pp. 611–615.
[14] A. Benjebbour, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, “System-level performance of downlink NOMA for future LTE enhancements,” in Proc. IEEE Globecom Workshops, June 2013, pp. 611–615.
[15] M. R. Hoiejj, J. Farah, C. A. Nour, and C. Douillard, “Resource allocation in downlink non-orthogonal multiple access (NOMA) for future radio access,” in Proc. IEEE VTC Spring, May 2015, pp. 1–6.
[16] P. Parida and S. S. Das, “Power allocation in OFDM based NOMA systems: A DC programming approach,” in Proc. IEEE Globecom Workshops, Dec 2014, pp. 1026–1031.
[17] N. Vucic, S. Shi, and M. Schubert, “DC programming approach for resource allocation in wireless networks,” in Proc. International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, May 2010, pp. 380–386.
[18] S. Hao, C. L., I. Z. Xu, and Q. Sun, “Energy efficiency and spectrum efficiency co-design: From NOMA to network NOMA,” IEEE Commun. Society MMTC E-Lett., vol. 9, no. 5, pp. 21–24, Sept 2013.
[19] Q. Sun, S. Han, C. L. I, and Z. Pan, “Energy efficiency optimization for fading MIMO non-orthogonal multiple access systems,” in Proc. IEEE ICC, June 2015, pp. 2668–2673.

[20] F. Fang, H. Zhang, J. Cheng, and V. C. M. Leung, “Energy-efficient resource allocation for downlink non-orthogonal multiple access network,” IEEE Trans. Commun., vol. 64, no. 9, pp. 3722–3732, Sept 2016.

[21] J. Huang, V. G. Subramanian, R. Agrawal, and R. Berry, “Joint scheduling and resource allocation in uplink OFDM systems for broadband wireless access networks,” IEEE J. Sel. A. Commun., vol. 27, no. 2, pp. 226–234, Feb 2009.

[22] R. Ruby and V. Leung, “Uplink scheduling solution for enhancing throughput and fairness in relayed long-term evolution networks,” IET Commun., vol. 6, no. 8, pp. 813–825, 2014.

[23] R. Ruby, V. C. M. Leung, and D. G. Michelson, “Uplink scheduler for SC-FDMA-based heterogeneous traffic networks with QoS assurance and guaranteed resource utilization,” IEEE Trans. Veh. Technol., vol. 64, no. 10, pp. 4780–4796, Oct 2015.

[24] M. Mollanoori and M. Ghaderi, “Uplink scheduling in wireless networks with successive interference cancellation,” IEEE Trans. Mobile Comput., vol. 13, no. 5, pp. 1132–1144, May 2014.

[25] K. Kumaran and L. Qian, “Scheduling on uplink of CDMA packet data network with successive interference cancellation,” in Proc. IEEE WCNC, vol. 3, March 2003, pp. 1645–1650 vol.3.

[26] M. Al-Imari, P. Xiao, and M. A. Imran, “Receiver and resource allocation optimization for uplink NOMA in 5G wireless networks,” in Proc. ISWCS, Aug 2015, pp. 151–155.

[27] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, “Iterative water-filling for gaussian vector multiple-access channels,” IEEE Trans. Inform. Theory, vol. 50, no. 1, pp. 145–152, Jan 2004.

[28] D. P. Bertsekas, “Nonlinear programming,” Athena Scientific, 1999.

[29] K. Hamidouche, W. Saad, and M. Debbah, “Many-to-Many matching games for proactive social-caching in wireless small cell networks,” in Proc. IEEE WiOpt, May 2014, pp. 569–574.

[30] A. Roth, “Stability and polarization of interests in job matching,” Econometrica, vol. 52, no. 1, pp. 47–57, 1984.

[31] M. Chiang, “Geometric programming for communication systems,” Commun. Inf. Theory, vol. 2, no. 1/2, pp. 1–154, Jul. 2005.

[32] D. Tse and P. Viswanath, “Fundamentals of wireless communication,” Cambridge University Press, 2005.

[33] K. McClaning, “Radio receiver design,” Noble Publishing Corporation, 2000.

[34] M. Hazewinkel, “Transmission rate of a channel,” Encyclopedia of Mathematics, Springer, 2001.

[35] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi, “A tutorial on geometric programming,” Optimization and Engineering, vol. 8, no. 1, p. 67, 2007.

[36] R. Ruby, V. C. M. Leung, and D. G. Michelson, “Centralized and game theoretical solutions of joint source and relay power allocation for AF relay based network,” IEEE Trans. Commun., vol. 63, no. 8, pp. 2848–2863, Aug 2015.

[37] X. Qiu and K. Chawla, “On the performance of adaptive modulation in cellular systems,” IEEE Trans. Commun., vol. 47, no. 6, pp. 884–895, Jun 1999.

[38] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatou, and M. Haenggi, “Superposition coding strategies: Design and experimental evaluation,” IEEE Trans. Wirel. Commun., vol. 11, no. 7, pp. 2629–2639, July 2012.

[39] L. Zhou, R. Ruby, H. Zhao, X. Ji, J. Wei, and V. Leung, “A graph-based resource allocation scheme with interference coordination in small cell networks,” in Proc. IEEE Globecom Workshops, 2014, pp. 211–218.