



# User Pairing for Downlink Non-Orthogonal Multiple Access Networks Using Matching Algorithm

DOI:

[10.1109/TCOMM.2017.2744640](https://doi.org/10.1109/TCOMM.2017.2744640)

## Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

## Citation for published version (APA):

Liang, W., Ding, Z., Li, Y., & Song, L. (2017). User Pairing for Downlink Non-Orthogonal Multiple Access Networks Using Matching Algorithm. *IEEE Transactions on Communications*. <https://doi.org/10.1109/TCOMM.2017.2744640>

## Published in:

IEEE Transactions on Communications

## Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

## General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [uml.scholarlycommunications@manchester.ac.uk](mailto:uml.scholarlycommunications@manchester.ac.uk) providing relevant details, so we can investigate your claim.



# User Pairing for Downlink Non-Orthogonal Multiple Access Networks using Matching Algorithm

Wei Liang, Zhiguo Ding, Yonghui Li and Lingyang Song

**Abstract**—In this paper, we study the user pairing in a downlink non-orthogonal multiple access (NOMA) network, where the base station allocates the power to the pairwise users within the cluster. In the considered NOMA network, a user with poor channel condition is paired with a user with good channel condition, when both their rate requirements are satisfied. Specifically, the quality of service for weak users can be guaranteed since the transmit power allocated to strong users is constrained following the concept of cognitive radio. A distributed matching algorithm is proposed in the downlink NOMA network, aiming to optimize the user pairing and power allocation between weak users and strong users, subject to the users' targeted rate requirements. Our results show that the proposed algorithm outperforms the conventional orthogonal multiple access scheme and approaches the performance of the centralized algorithm, despite its low complexity. In order to improve the system's throughput, we design a practical adaptive turbo trellis coded modulation scheme for the considered network, which adaptively adjusts the code rate and the modulation mode based on the instantaneous channel conditions. The joint design work leads to significant mutual benefits for all the users as well as the improved system throughput.

**Index Terms**—Non-orthogonal multiple access scheme, Matching Theory, Adaptive Trellis Turbo Coded Modulation.

## I. INTRODUCTION

The non-orthogonal multiple access (NOMA) technique emerged as a promising multiple access (MA) technique to improve the access efficiency of the future mobile network [1]–[3]. Cooperative communication in conjunction with NOMA scheme was characterized in [4], in which a cooperative NOMA scheme was conceived for exploiting the fact that the user with better channel condition have the prior information about the message for other users. The concept of NOMA have been applied in multiple research areas, such as the multiple-input multiple-output (MIMO) systems [5], high-rate visible light communication (VLC) downlink networks [6], simultaneous wireless information and power transfer (SWIPT) [7] as well as the physical layer security [8]. The key idea of the NOMA technique is to explore the power domain for multiple access (MA). More specifically, in a downlink NOMA scheme, a base station (BS) can serve multiple users within the same time/frequency channel via different power allocation

The work was supported in part by the UK EPSRC under grant number EP/L025272/1, in part by H2020-MSCA-RISE-2015 under grant number 690750, in part by National Science and Technology Major Project of China under grant number 2016ZX03001017, and in part by ARC under grant number DP150104019.

W. Liang and Z. G. Ding are with school of Computing and Communications, Lancaster University, LA1 4WA, UK. Email: {w.liang, z.ding}@lancaster.ac.uk.

Y. H. Li is with school of Electrical and Information Engineering, University of Sydney, NSW, 2006, Australia. Email: yonghui.li@sydney.edu.au.

L. Y. Song is with school of Electrical Engineering and Computer Science, Peking University, 100871, China. Email: lingyang.song@pku.edu.cn.

coefficients, where the users with poorer channel conditions are given more transmission power [4], [9]. The users with more power allocated are capable of decoding their own messages by treating the other users' information as noise. On the other hand, the users with less power allocated will employ the successive interference cancellation (SIC) technique to decode their own messages by first removing the other users' information from their observations [5], [9]. A cognitive radio inspired NOMA networks has been studied from [5], [10], [11], which can be termed as CR-NOMA. The concept of power allocation in CR-NOMA network was proposed in [5], has investigated a more effective approach to allocate power in user pairing-based NOMA systems. In the CR-NOMA network [5], [11], a user viewed as a cognitive user (CU) with strong channel condition accesses the spectrum occupied by a user seen as a primary user (PU) with poor channel condition under a predefined interference constraints. Most of the existing works of the CR-NOMA systems focused on performance analysis, such as the outage probability [5], [12]. There are very limited studies on the user pair problem in CR-NOMA. Nevertheless, motivating to maximize the system's throughput, the user pairing in CR-NOMA systems leads to combinatorial problem, and is very challenging, thereby becoming NP-hard. In this paper, we employ the matching theory to solve the user pair problem in CR-NOMA network, for the sake of finding promising solutions with relatively small complexity.

Matching theory [13], [14] has been known as an efficient technique to solve the combinatorial problem of matching players in two distinct sets, by using the players' individual information and their preferences. The two-side stable matching problem has been widely investigated from both theoretical and practical perspectives in [13]–[18]. The most classical matching problems are one-to-one matching [17], many-to-one matching [16] and many-to-many matching [18]. In the one-to-one matching problem, each player can be matched to at most one member from the opposite set. In contrast to one-to-one matching, at least one player in the many-to-one matching problem can be matched to multiple players in the opposing set. In the context of many-to-many matching, at least one player in each of the two sets can be matched to more than one member from the other set. In [18], a many-to-many two sided user-subchannel matching algorithm has been proposed for downlink NOMA networks, aim to optimize the sub-channel assignment and power allocation problem. In the existing literature, there is no study devoted to implement the matching theory on the considered CR-NOMA systems.

In this paper, we investigate the user pairing in the CR inspired NOMA networks, where two users within a user pair

are allowed to simultaneously transmit information. Specifically, the paired two users share the same spectrum resources with the aid of allocating different power levels. The motivation of our considered CR-NOMA is to resist the spectral efficiency loss when users' channel conditions are poor, and these weak users<sup>1</sup> can be served for the sake of improving the user fairness. Specifically, the users with strong channel condition would have more chances to access the weak users' spectrum bands which will not cause too much performance degradation of the weak users. Different to [18], we associate our user pairing problem with equivalent to two-side one-to-one matching problem, which aims to increase the individual data rates of the paired users as well as the throughput of the overall network, while both the minimum rate requirements of the PUs and CUs are satisfied. In [17], [18], the time slots, frequency bands and sub-channel assignments are the exchange proposals. In contrast to the implementations of the matching theory in NOMA and CR systems, in our paper, the CUs negotiate with the PUs regarding the amount of power coefficients available for them, then these CUs are allowed to access their paired PUs' spectrum bands. Additionally, we also develop a more realistic system, wherein considering the imperfect channel state information (CSI). Finally, we design the practical adaptive coding and modulation schemes based on turbo trellis coded modulation (TTCM) for our CR-NOMA system.

The key contributions of the paper are summarized as follows:

1) We conceive a downlink CR-NOMA network, in which the CUs could transmit simultaneously with the PUs under the quality of service (QoS) constraint that the interferences inflicted by the CUs on the PUs do not degrade the PUs' communication quality. The PUs and CUs are assumed to be paired with each other in the CR-NOMA network. Particularly, the PU and CU, with two very different channel conditions, will be paired in each small cluster in order to perform the NOMA technique in the power domain, while reducing the system load. Additionally, the achievable rate of the CUs can be increased by employing the SIC technique, which leads to an improved overall system throughput as well.

2) We propose a distributed algorithm for the user pairing and power allocation in the CR-NOMA system based on the one-to-one matching theory, where the PUs negotiate with the CUs for obtaining the appropriate power allocations, aiming to maximize the system's throughput. For the sake of comparison, we develop two extreme benchmarks. The first one with the highest complexity demanded, namely centralized algorithm, is to find the optimal pairing of the PUs and CUs by exhaustive search, while assigning transmit power optimally. Another called random algorithm, is to randomly pair the PUs and CUs and allocate power in a fixed manner.

3) To address practical issues, we investigate the impact of imperfect CSI on the performance of the CR-NOMA network. More specifically, we assume the channel estimation error model, where the BS has an estimation of the channel and

<sup>1</sup>In the CR inspired NOMA networks, the weak user can be viewed as the primary user (PU) and the strong user can be seen as the cognitive user (CU) [5], [11].

a priori knowledge of the variance of the estimation error. Additionally, we also design a more practical CR-NOMA network, which is facilitated by the bandwidth-efficient ATTCM arrangement in order to improve the system's throughput performance. The transmission rate/throughput of the system is adapted according to the instantaneous channel conditions. 4) We provide a comprehensive performance analysis for our CR-NOMA system with/without the practical ATTCM scheme, when employing three user pairing and power allocation algorithms. Our simulation results show that, in comparison with the OMA assisted CR system, our CR-NOMA system can always achieve a significant throughput enhancement, regardless of user pairing and power allocation algorithms employed. Additionally, our proposed conventional distributed algorithm significantly outperforms the random algorithm, and its throughput performance is close to the centralized algorithm's. However, when considering imperfect CSI, the throughput performance of our proposed DMA gets much closer to that of the CA, which reveals that our distributed algorithm is a promising solution to practical CR-NOMA systems. Importantly, we also show that, the CR-NOMA system's throughput can be further improved at a slight cost of using the ATTCM scheme.

The rest of the paper is organized as follows. The system model of the CR-NOMA network is outlined in Section II. The general problem of the user pairing and power allocation is discussed in Section III. The proposed distributed approach for solving the user pairing problem is introduced in Section IV. A practical adaptive coded modulation scheme is described in Section VI. The theoretical performance of the proposed CR-NOMA scheme is evaluated in Section V and the practical performance of the ATTCM aided CR-NOMA scheme is investigated in Section VI-B. Finally, our conclusions are presented in Section VII.

## II. SYSTEM MODEL

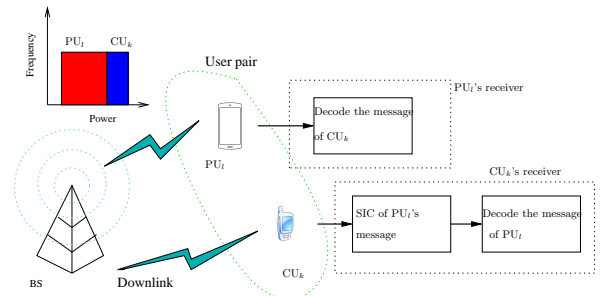


Fig. 1. The system model of user pairing among PU<sub>l</sub> and CU<sub>k</sub> in CR-NOMA networks.

Refer to Fig. 1, each cluster in our CR-NOMA system consists of a single BS, as well as the user  $l$  and user  $k$ . As shown in Fig. 2, the users are assumed to be uniformly deployed in a disc, namely  $\mathcal{D}_o$  with the radius  $r_o$ . The BS is located at the center of this disc. We further assume that this disc is divided into two regions. A smaller disc  $\mathcal{D}_i$  with radius  $r_i$  is located inside the disc  $\mathcal{D}_o$  and they have the same origin. Specifically, the radius of  $\mathcal{D}_i$  is smaller than that of the  $\mathcal{D}_o$ , which is  $r_o > r_i$ . Additionally,  $K$  users are assumed to be randomly distributed in  $\mathcal{D}_i$ . On the contrary,

$L$  users are uniformly distributed in a region, in which  $\mathcal{D}_i$  is subtracted from  $\mathcal{D}_o$ , i.e.  $|\mathcal{D}_o - \mathcal{D}_i|$ . Hence, user  $l$  and user  $k$  are randomly scheduled and matched with each other. In the considered downlink CR inspired NOMA system, the  $L$  users who located at the edge of the cell are treated as PUs. Specifically, those stronger users which located close to the

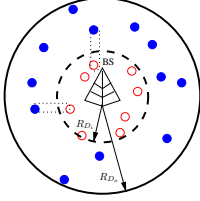


Fig. 2. The system design of distance setting.

BS can be viewed as the CUs. For the sake of improving the spectrum efficiency, those CUs have been squeezed into the spectrum occupied by the PUs. Therefore, the so-called PUs and CUs have the significant channel condition difference in our considered system. Without loss of generality, the channels are sorted as  $0 < |h_l^p|^2 < \dots < |h_L^p|^2 < |h_k^c|^2 < \dots < |h_K^c|^2$ , where  $h_l^p$  defines the channel from BS to the  $l$ th PU and  $h_k^c$  is the channel from the BS to the  $k$ th CU. We assume that the CUs and PUs can transmit simultaneously, and each PU can be paired with one particular CU. The paired PU and CU can transmit information on the same spectrum, but users from different pairs should transmit on orthogonal channels. Additionally, the pathloss is defined as  $\rho = 1/d_{ab}^\alpha$  [19], where  $d_{ab}$  is the geometrical distance between node  $a$  and node  $b$ , and the path-loss exponent considered in our system is  $\alpha = 4$ . Similar to [20], [21], all channels in our system are assumed to impose quasi-static Rayleigh fading, where the channel coefficients are constant for each transmission block, but vary independently between different blocks and all the CSI of the users are assumed to be known at BS.

The signal received at PU $_l$  is given by

$$y_l^{PU} = h_l^p(\sqrt{\alpha_l^p P_S} x_l^p + \sqrt{\alpha_k^c P_S} x_k^c) + n_l^p, \quad (1)$$

and the signal received at CU $_k$  is expressed as:

$$y_k^{CU} = h_k^c(\sqrt{\alpha_l^p P_S} x_l^p + \sqrt{\alpha_k^c P_S} x_k^c) + n_k^c. \quad (2)$$

where  $P_S$  is the transmit power emitted from the BS. The term  $\sqrt{\alpha_l^p P_S} x_l^p + \sqrt{\alpha_k^c P_S} x_k^c$  of Eq. (1) and Eq. (2) represents the combined signal of PU $_l$  and CU $_k$  emitted from the BS. The channels  $h_l^p$  and  $h_k^c$  are assumed to be Rayleigh distributed, following the complex-valued Gaussian distribution of  $\mathcal{CN}(0, \sigma^2)$ . In Eq. (1) and Eq. (2),  $n_l^p$  and  $n_k^c$  are the Gaussian noise terms, which have a zero mean and a noise variance of  $N_0/2$  per dimension. Additionally,  $\alpha_l^p$  and  $\alpha_k^c$  represent the power coefficients for PU $_l$  and CU $_k$ , and we have  $\alpha_k^c + \alpha_l^p = 1$ . For simplify, we use the notation  $\alpha_{l,k}$  to represent the power allocation coefficient, and we define  $\alpha_{l,k} = \alpha_l^p$ . Then, the power allocated to the CU is  $(1 - \alpha_{l,k})$ . The SIC technique<sup>2</sup> is only employed at the CU within each

<sup>2</sup>The SIC is applied at the user device in the considered CR inspired NOMA system, in order to cancel the multi-user interference. Therefore, in our proposed framework, only two users within each pair share the same spectrum. Additionally, the SIC is only employed by the user with strong channel condition. Hence, the complexity of implementing SIC is just one and is affordable at user device.

pair. Therefore, the CU $_k$  would detect the PU $_l$ 's message when  $h_l^p < h_k^c$ , and then remove the message from its observation in a successive manner. Meanwhile, the message of the CU $_k$  will be treated as noise at PU $_l$ .

Then the achievable rate of PU $_l$  is given by [5], [9]:

$$R_l^{PU} = \log_2 \left( 1 + \frac{\gamma_l^{PU} |h_l^p|^2 \alpha_{l,k}}{1 + \gamma_l^{PU} |h_l^p|^2 (1 - \alpha_{l,k})} \right), \quad (3)$$

Note that the transmit signal to noise ratio (SNR)<sup>3</sup> of the  $l$ th PU is  $\gamma_l^{PU} = \frac{\alpha_l^p P_S}{N_0}$  and that of the  $k$ th CU is  $\gamma_k^{CU} = \frac{\alpha_k^c P_S}{N_0}$ . As a result, the achievable rate of CU $_k$  can be written as:

$$R_k^{CU} = \log_2 (1 + (1 - \alpha_{l,k}) \gamma_k^{CU} |h_k^c|^2), \quad (4)$$

Specifically, the PU is willing to perform NOMA with the CU only if it can achieve a higher rate in comparison with the conventional orthogonal multiple access (OMA) transmission [5].

We set the rate of the PU $_l$  using conventional OMA-CR as the minimum rate requirement, namely  $R_{l,req}^{PU}$ , and it can be expressed as:

$$R_{l,req}^{PU} = \frac{1}{2} \log_2 (1 + \gamma_l^{PU} |h_l^p|^2), \quad (5)$$

where the factor  $\frac{1}{2}$  of Eq.(5) is due to the fact that conventional OMA results in a multiplexing loss of  $\frac{1}{2}$ . Specifically, PUs and CUs are served in a TDMA manner for conventional OMA-CR which has a multiplexing loss of  $\frac{1}{2}$  for the PUs. Additionally, the minimum rate constraints of PU $_l$  is defined as  $R_l^{PU} \geq R_{l,req}^{PU}$ , which can be re-written as:

$$\log_2 \left( 1 + \frac{\gamma_l^{PU} |h_l^p|^2 \alpha_{l,k}}{1 + \gamma_l^{PU} |h_l^p|^2 (1 - \alpha_{l,k})} \right) \geq \frac{1}{2} \log_2 (1 + \gamma_l^{PU} |h_l^p|^2). \quad (6)$$

After that, the power allocation fraction of CU $_k$  can be derived as:

$$(1 - \alpha_{l,k}) \leq \frac{\sqrt{1 + \gamma_l^{PU} |h_l^p|^2} - 1}{\gamma_l^{PU} |h_l^p|^2}. \quad (7)$$

Moreover, we have a minimum rate requirement of the CU $_k$ , namely  $R_{k,req}^{CU}$ , in order to make sure CU $_k$  can benefit from accessing the PU's spectrum band. Furthermore, the minimum rate requirement of each CU is dependent on its channel conditions.

### III. FORMULATION OF THE OPTIMIZATION PROBLEM

Let us define a  $(L \times K)$ -element matching matrix  $\mathbf{M}$ . When PU $_l$  is matched to CU $_k$ , their corresponding matching matrix entry is given by  $m_{l,k} = 1$ , otherwise  $m_{l,k} = 0$ . We also define a  $(L \times K)$ -element power allocation matrix  $\mathbf{P}$  with elements  $\alpha_{l,k}$ . In particular, when given the users' predefined QoS requirements, the objective function of maximizing the system throughput are formulated as:

$$\{\mathbf{M}^{cen}, \mathbf{P}^{cen}\} = \max_{\mathbf{M}, \mathbf{P}} \sum_{l=1}^L \sum_{k=1}^K m_{l,k} W(\alpha_{l,k}), \quad (8)$$

<sup>3</sup>The definition of the transmit SNR is unconventional, because it relates the transmit-power to the noise-power at the receiver, which are quantities measured at different locations. Nonetheless, this convenient definition simplifies our discussion, as proposed in [22].

$$\begin{aligned}
s.t. \quad & (a) R_l^{PU}(\alpha_{l,k}) \geq R_{l,req}^{PU}, \forall l \forall k \\
& (b) R_k^{CU}(1 - \alpha_{l,k}) \geq R_{k,req}^{CU}, \forall l \forall k \\
& (c) 0 \leq \alpha_{l,k} \leq 1, \forall l \forall k \\
& (d) \sum_{l=1}^L m_{l,k} \leq 1, \forall k \\
& (e) \sum_{k=1}^K m_{l,k} \leq 1, \forall l \\
& (f) m_{l,k} \in \{0, 1\}, \forall l \forall k \\
& (g) N_f = \sum_{l=1}^L \sum_{k=1}^K m_{l,k}, \forall l \forall k
\end{aligned}$$

The conditions (a) and (b) ensure that the minimum rate requirements of the PUs and CUs can be achieved. The condition (c) specifies the range of the power allocation factor of the PU. Therefore, conditions (d) and (e) ensure that each PU (or CU) will only be matched to one CU (or PU). The condition (f) states that the corresponding matching matrix entry is either 1 or 0. Conditions (g) defines the total number of available frequency bands  $N_f$  is equal to the number of PUs. Without loss of generality, we assume that each PU occupies a single spectrum band. Moreover, the choice of the weight  $W(\alpha_{l,k})$  of Eq. (8) depends on the objective function to be optimized. We consider three cases to set up the relative optimization functions, which are described below:

- Case 1: objective function is to maximize the sum-rate of matched PUs:

$$W_{PU}(\alpha_{l,k}) = R_l^{PU}(\alpha_{l,k}). \quad (9)$$

- Case 2: objective function is to maximize the sum-rate of matched CUs:

$$W_{CU}(\alpha_{l,k}) = R_k^{CU}(1 - \alpha_{l,k}). \quad (10)$$

- Case 3: objective function is to maximize the total sum-rate of matched PUs and CUs:

$$W_{total}(\alpha_{l,k}) = R_l^{PU}(\alpha_{l,k}) + R_k^{CU}(1 - \alpha_{l,k}). \quad (11)$$

Note that, the objective function of Eq. (8) is a function of continuous variables  $\alpha_{l,k}$  and binary variables  $m_{l,k}$ . The objective function in Eq. (8) can be decoupled with respect to two entries, which are the optimal power allocation matrix  $\mathbf{P}$  and the optimal matching matrix  $\mathbf{M}$  [23]. Hence we solve this optimization problem in two steps in Section III-A.

#### A. Decoupling the Optimization Function

For a given matched pair (PU, CU), mapping is valid only if the conditions (a) and (b) of Eq. (8) have been satisfied. During the first step, the optimum value,  $\alpha_{l,k}^{cen}$ , and its corresponding weight  $W(\alpha_{l,k}^{cen})$  for each pair (PU, CU) can be found under the conditions (a), (b) and (c) of Eq. (8), where the computations of the achievable rate of the PU and the CU have been introduced in Section II. Note that all  $W(\alpha_{l,k})$  depend on a single variable  $\alpha_{l,k}$  and then maximizing the individual  $W(\alpha_{l,k})$  leads to the maximal value of the overall objective. Suppose we have a particular matched pair (PU<sub>*i*</sub>, CU<sub>*j*</sub>), where

their corresponding matching matrix entry can be defined as  $m_{i,j}$ . Then the optimization value of  $\alpha_{l,k}$  can be formulated as:

$$\alpha_{l,k}^{cen} = \max_{\alpha_{l,k}} W(\alpha_{l,k}) m_{i,j}, \quad (12)$$

where  $m_{i,j}$  satisfies the matching constraints in (d), (e) and (f) of Eq. (8).

According to condition (a) of Eq. (8), the lower bound of  $\alpha_{l,k}$  is given by

$$\alpha_{l,k}^{\min} \geq \frac{\gamma_l^{PU} |h_l^p|^2 + 1 - \sqrt{\gamma_l^{PU} |h_l^p|^2 + 1}}{\gamma_l^{PU} |h_l^p|^2}. \quad (13)$$

Moreover, from condition (b) of Eq. (8) the upper bound of  $\alpha_{l,k}$  can be given by

$$\alpha_{l,k}^{\max} \leq \frac{1 + \gamma_k^{CU} |h_k^c|^2 - 2R_{k,req}^{CU}}{\gamma_k^{CU} |h_k^c|^2}. \quad (14)$$

By considering the three cases of the optimization problems as described in Eq. (9), Eq. (10) and Eq. (11), the setting of  $\alpha_{l,k}^{cen}$  in Eq. (12) are derived as:

- Case 1:

$$\alpha_{l,k}^{cen} = \frac{1 + \gamma_k^{CU} |h_k^c|^2 - 2R_{k,req}^{CU}}{\gamma_k^{CU} |h_k^c|^2}. \quad (15)$$

- Case 2:

$$\alpha_{l,k}^{cen} = \frac{\gamma_l^{PU} |h_l^p|^2 + 1 - \sqrt{\gamma_l^{PU} |h_l^p|^2 + 1}}{\gamma_l^{PU} |h_l^p|^2}. \quad (16)$$

- Case 3:

$$\alpha_{l,k}^{cen} = \begin{cases} \frac{1 + \gamma_k^{CU} |h_k^c|^2 - 2R_{k,req}^{CU}}{\gamma_k^{CU} |h_k^c|^2} & C_{CU}^* < C_{PU}^*; \\ \frac{\gamma_l^{PU} |h_l^p|^2 + 1 - \sqrt{\gamma_l^{PU} |h_l^p|^2 + 1}}{\gamma_l^{PU} |h_l^p|^2} & C_{CU}^* > C_{PU}^*; \end{cases} \quad (17)$$

The derivation of Eq. (17) is detailed in Appendix A. After obtaining the optimization value of  $\alpha_{l,k}^{cen}$  of Eq. (12), the optimization function of Eq. (8) can be re-written as :

$$\{\mathbf{M}^{cen}\} = \max_M \sum_{l=1}^L \sum_{k=1}^K W(\alpha_{l,k}^{cen}) m_{l,k}, \quad (18)$$

$$\begin{aligned}
s.t. \quad & (d) \sum_{l=1}^L m_{l,k} \leq 1, \forall k \\
& (e) \sum_{k=1}^K m_{l,k} \leq 1, \forall l, \\
& (f) m_{l,k} \in \{0, 1\} \quad \forall l \forall k.
\end{aligned}$$

The weight  $W(\alpha_{l,k}^{cen})$  can be solved according to the terms  $R_l^{PU}(\alpha_{l,k}^{cen})$ ,  $R_{l,req}^{PU}$ ,  $R_k^{CU}(\alpha_{l,k}^{cen})$  and  $R_{k,req}^{CU}$ . These four terms can be calculated based on the known channel conditions. This follows the fact that the weight  $W(\alpha_{l,k}^{cen})$  only depends on the variable  $\alpha_{l,k}^{cen}$ . Specifically,  $\alpha_{l,k}^{cen}$  is the maximum value of variable  $\alpha_{l,k}$ . In a given matching pair,  $\alpha_{l,k}^{cen}$  can be obtained from Eq (15), Eq. (16) and Eq. (17) according to different optimization objectives. Note that, the optimal objective of

Eq. (18) is  $W(\alpha_{l,k}^{cen})m_{l,k}$ , which is a specific example of the formulation shown in Eq. (8), where  $W(\alpha_{l,k})$  of Eq. (8) is substituted by  $W(\alpha_{l,k}^{cen})$  of Eq. (18). Therefore, the solution for Eq. (18) is a feasible solution for Eq. (8), which is equivalent to find the best matching pair.

### B. Centralized Algorithm

In order to find the optimal solution, we use the centralized algorithm (CA). We consider all possible matching pairs of the PUs and CUs, and then select that particular matched pair, which has the maximum sum-rate according to our intended objective functions as shown in Eq. (9), Eq. (10) and Eq. (11). By referring to Eq. (18), the centralized solution requires an exhaustive search over all possible matching pairs and the power allocation combinations. Specifically, the CA relies on an exhaustive search method that imposes the highest number of operations for the sake of finding the optimum solution. Therefore, the CA has the highest complexity. Additionally, The amount of complexity required for this CA increases with the number of users, which may become quite high, rendering it impractical. Therefore, we will propose a distributed approach in association with the matching algorithm as described in Section IV.

## IV. THE DISTRIBUTED APPROACH

### A. Preference Lists

Our problem can be modeled as a two-side one-to-one matching problem [13], [14], in which a set of PUs will be matched with a set of CUs. We consider a canonical matching problem [15], where the preferences of one player set depends on the other player set. Specifically, before the information exchange (proposals), each player is required to merely collect the information from the opposite set users they are interested, and performs a ranking according to its preferences. In the considered CR-NOMA system, the BS broadcasts the ranking information of the channels to all the PUs and CUs. According to the objective functions presented in Eq.(9), Eq.(10) and Eq.(11), we treat one sets of users as the proposers and the opposite sets of users as selectors. The users who to be the proposers or/and selectors depend on the consideration of the objective functions. In case 1 of Eq.(9), PUs are act as the proposers and CUs are the selectors, then before any offer is made to the CUs, each of the PUs construct a preference list of CUs, which can satisfy the PU's rate requirement. Thus, the preference list for  $PU_l$  is given by:

$$PULIST_l = \{CU_{\kappa(k)}\}_{k=1}^{K'} , \quad (19)$$

where the function  $\kappa(k)$  satisfies the following condition:

$$R_{l,\kappa(k)}^{PU}(\alpha_{l,\kappa(k)}) > R_{l,req}^{PU}, \quad k \in (1, \dots, K) . \quad (20)$$

The indexes of the CUs are recorded in the PULIST, when their corresponding rate have satisfied the PU's rate requirement.  $K'$  in Eq. (19) is the number of CUs in  $PULIST_l$ , where  $K' \leq K$ . Additionally, we have assumed that the first  $CU_{\kappa(k)}$  at the top of the  $PULIST_l$  provides the highest rate  $R_{l,\kappa(k)}^{PU}(\alpha_{l,\kappa(k)})$ . Similarly, each CU also has its preferred PU

list. When a CU transmits on the spectrum band occupied by the preferred PU, its achievable transmission rate should be higher than its minimum rate requirement  $R_{k,req}^{CU}$ . Thus, the preference list for  $CU_k$  is given by:

$$CULIST_k = \{PU_{\iota(l)}\}_{l=1}^{L'} , \quad (21)$$

where the function  $\iota(l)$  satisfies the following condition:

$$R_{\iota(l),k}^{CU}(\alpha_{\iota(l),k}) > R_{k,req}^{CU}, \quad l \in (1, \dots, L) . \quad (22)$$

Again, the elements in  $CULIST_k$  are ranked in decreasing order in terms of the achievable rate.  $L'$  in Eq. (21) is the number of PUs in  $CULIST_k$ , i.e.  $L' \leq L$ . Additionally, the PUs and CUs would exchange their roles in the Case 2 scenario. Specifically, the CUs would make offers to the PUs in the Case 2 scenario. In the Case 3, the PUs and CUs would compare first, and then decide which one to be the proposers. This is based on the optimization results of Eq. (11).

### B. Conventional Distributed Matching Algorithm (DMA)

Different to the CA as described in Section III, we motivate to use the matching theory to find a suboptimal solution of Eq. (8). In the context of the CA, the BS needs to know all the CSIs of all the PUs and CUs. Additionally, in our proposed DMA, there is no central controller and does not require the BS to know the full CSI of users, which reduces the system complexity. Instead, the BS only needs to have the rank information of channels, and is able to broadcast this information to PUs and CUs for operating distributed user pairing and power allocation. Then, PUs and CUs can construct their preference list without any CSI. The key idea of this distributed algorithm in Case 1 scenario is that each PU make an offer to its most preferred CU from its preference list. Then each CU has right to accept or reject these offers. When all PUs make their offer once, one round of proposals is preformed. In order to construct the user pairs among the PUs and CUs in the considered CR inspired NOMA networks, the PU having a poorer channel condition matches with the CU having a better channel condition. Then we distinguish the matched PUs and CUs in the power domain, aiming to improve the total sum-rate of the PUs and CUs. Additionally, each CU and its paired PU transmit simultaneously in appropriate different power levels by exchanging their information. The BS allocates the initialized power allocations to the PU and CU, which the PU would get more power than the CU. After that, the PUs and CUs negotiate with each other for a specific value of power, i.e.  $\alpha_{l,k}$ , as introduced in Section III. Employing the DMA for efficiently representing the interaction among the PUs, where each PU chooses its allocation independently of the others.

Additionally, we first initialize the power allocation factor of  $PU_l$  as  $\alpha_{l,k} = \alpha_{init}$  and set the value of the power step counter to  $\tau$ . We construct  $PULIST_l$  and  $CULIST_k$  based on the initialized value  $\alpha_{init}$ . As an example, an offer of  $\alpha_{l,k}$  is given by  $PU_l$  to the  $CU_k$  in the top of its preference list  $PULIST_l$ . Then this  $CU_k$  grants a power allocation factor  $1 - \alpha_{l,k}$  for itself. When  $CU_k$  receives an offer, it has two options: (1) reject the offer if  $PU_l$  is not in  $CU_k$ 's  $CULIST_k$ ; (2) accept

TABLE I

CONVENTIONAL DISTRIBUTED MATCHING ALGORITHM (DMA) IN CASE

1

---

**Step 1: Initialization**

- 1) Set  $\alpha_{l,k} = \alpha_{init}$ , and set  $\tau$ .
- 2) Construct  $PULIST_l$  based on  $\alpha_{init}$ ,  $l \in L$ .
- 3) Construct  $PU_{NOTMATCH}$  to list the PUs which are not matched, yielding  $PU_{NOTMATCH} = \{PU_1, \dots, PU_L\}_{l=1}^L$ .
- 4) Construct  $CU_{NOTMATCH}$  to list the CUs which are not matched, yielding  $CU_{NOTMATCH} = \{CU_1, \dots, CU_K\}_{l=1}^K$ .
- 5) Construct a list of all PUs which are matched, yielding  $PU_{MATCH} = \{PU_{MATCH}(l)\}_L$ , where  $PU_{MATCH}(l) = 1$  if the PU is matched to the CU, otherwise set  $PU_{MATCH}(l) = 0$ .
- 6) Set  $l = 1$  and  $k = 1$ .

**Step 2: The PUs negotiate with the CUs:**

- Step 2A: PU selects its current best CU:  $PU_l$  offers  $\alpha_{l,k} = \alpha_{init}$  to the top CU from its  $PULIST_l$ .
  - 1) If this  $PU_l$  is in  $CULIST_k$  of  $CU_k$ :
    - $PU_l$  and  $CU_k$  are temporal matched.
    - Remove  $PU_l$  from  $PU_{NOTMATCH}$ . Update the  $l$ th element of  $PU_{MATCH}$  to  $PU_{MATCH}(l) = 1$ . Remove  $CU_k$  from  $CU_{NOTMATCH}$ .
  - 2) Else:
    - Update the value of power as  $\alpha_{l,k}^* = \alpha_{l,k} - \tau$ .
    - Update  $PULIST_l$  based on  $\alpha_{l,k}^*$ , then go back to Step 2A and set  $PU_{MATCH}(l) = 0$ . Remove  $PU_l$  from the set  $PU_{NOTMATCH}$ , if  $PULIST_l = \emptyset$  and go to Step 3.
- Step 2B: If this preferred  $CU_k$  has a current matched pair,  $PU_{cur}$ .
  - 1) if  $(1 - \alpha_{l,k}) > (1 - \alpha_{l_{cur},k})$ , this  $CU_k$  would discard its current matching  $PU_{cur}$ . Then  $PU_l$  and  $CU_k$  are matched and the  $PU_{cur}$  goes to Step 2A-(2).
  - 2) Else: The  $PU_l$  goes to Step 2A-(2).

**Step 3: If no further matching is required**  
Then go to Step 4, otherwise go to Step 2 for the next PU,  $l = l + 1$ .

**Step 4: End of the the algorithm.**

---

TABLE II

THE ALGORITHM FOR CASE 2 SCENARIO

---

**Step 1: Initialization**

- 1) Set  $\alpha_{l,k} = \alpha_{init}$ , and set  $\tau$ .
- 2) Construct  $CULIST_k$  based on  $(1 - \alpha_{init})$ ,  $k \in K$ .
- 3) Construct  $PU_{NOTMATCH}$  to list the PUs which are not matched, yielding  $PU_{NOTMATCH} = \{PU_1, \dots, PU_L\}_{l=1}^L$ .
- 4) Construct  $CU_{NOTMATCH}$  to list the CUs which are not matched, yielding  $CU_{NOTMATCH} = \{CU_1, \dots, CU_K\}_{l=1}^K$ .
- 5) Construct a list of all CUs which are matched, yielding  $CU_{MATCH} = \{CU_{MATCH}(k)\}_K$ , where  $CU_{MATCH}(k) = 1$  if the CU is matched to the PU, otherwise set  $CU_{MATCH}(k) = 0$ .
- 6) Set  $k = 1$  and  $l = 1$ .

**Step 2: The CUs negotiate with the PUs:**

- Step 2A: CU selects its current best PU:  $CU_k$  offers  $1 - \alpha_{l,k} = \alpha_{init}$  to the top PU from its  $CULIST_k$ .
  - 1) If this  $CU_k$  is in  $PULIST_l$  of  $PU_l$ :
    - $CU_k$  and  $PU_l$  are temporal matched.
    - Remove  $CU_k$  from  $CU_{NOTMATCH}$ . Update the  $k$ th element of  $CU_{MATCH}$  to  $CU_{MATCH}(k) = 1$ . Remove  $PU_l$  from  $PU_{NOTMATCH}$ .
  - 2) Else:
    - Update the value of power as  $\alpha_{l,k}^* = \alpha_{l,k} - \tau$ .
    - Update  $CULIST_k$  based on  $(1 - \alpha_{l,k}^*)$ , then go back to Step 2A and set  $CU_{MATCH}(k) = 0$ . Remove  $CU_k$  from the set  $CU_{NOTMATCH}$ , if  $CULIST_k = \emptyset$  and go to Step 3.
- Step 2B: If this preferred  $PU_l$  has a current matched pair,  $CU_{cur}$ .
  - 1) if  $\alpha_{l,k} > \alpha_{l_{cur},k}$ , this  $PU_l$  would discard its current matching  $CU_{cur}$ . Then  $CU_k$  and  $PU_l$  are matched and the  $CU_{cur}$  goes to Step 2A-(2).
  - 2) Else: The  $CU_k$  goes to Step 2A-(2).

**Step 3: If no further matching is required**  
Then go to Step 4, otherwise go to Step 2 for the next CU,  $k = k + 1$ .

**Step 4: End of the the algorithm.**

---

it, when this  $PU_l$  is in the  $CULIST_k$ , creating a matching pair between  $PU_l$  and  $CU_k$ . Additionally, if two or more PUs send offers to  $CU_k$ , a conflict will happen [16]. In order to avoid this conflict,  $CU_k$  will only pair with the PU who could provide a higher revenue. Specifically, if this intended  $CU_k$  has already been matched to  $PU_{cur}$ , and this  $PU_{cur}$  failed to provide a lower power allocation factor to the  $CU_k$ , i.e.

$$\alpha_{l_{cur},k} > \alpha_{l,k}, \quad (23)$$

then  $CU_k$  discards its current matching in favor of the new matching. This is because a larger value of  $\alpha_{l,k}$  leads to a lower achievable rate of the CU, which is refer to Eq. (4). Moreover, the rejected  $PU_{cur}$  will update the value of the power allocated factor by setting it to

$$\alpha_{l_{cur},k}^* = \alpha_{l_{cur},k} - \tau. \quad (24)$$

and then it will reconstruct its preference list based on  $\alpha_{l_{cur},k}^*$  and repeat the matching process. This algorithm aims to find the specific power allocation factors, which can be accepted by both the  $PU_l$  as well as the  $CU_k$ , and the algorithm is terminated when each PU has found its appropriate matched partner, provided that both their rate requirements can be satisfied. Similar to the DMA of Case 1 as shown in Table I, when we consider the Case 2 scenario, the CUs become the proposer and the PUs are the selectors. The details are

described in Table II. Additionally, when we consider the Case 3 scenario, the PUs and CUs would take turn to be the proposers depending on their current avenue as shown in Table III. Note that, we have some particular situations that two players may compete with each other for appealing the same choice. During the competition, each player reduces its own power allocation factor, and this competition completes when one of the user's achievable rate is inferior to its minimum rate requirement. Then we allow the losing user to select its surplus best choice from its current preference list.

## V. PERFORMANCE ANALYSIS OF THEORETICAL CR-NOMA NETWORK

### A. The Stability of our Conventional DMA

Before introducing the definition of stable matching, we would like to highlight our related notations and definitions. We denote the set of PUs' indexes by  $\mathcal{L} = \{1, \dots, l, \dots, L\}$  and the set of CUs' indexes by  $\mathcal{K} = \{1, \dots, k, \dots, K\}$ .

**Definition 1:** We define a matching  $\mu$  as a one-to-one matching of the set  $\mathcal{L} \cup \mathcal{K}$ . We refer to  $\mu(l)$  and  $\mu(k)$  as the partner of  $l$  and  $k$ , respectively. Then we have  $\mu(l) \in \mathcal{K}$  and  $\mu(k) \in \mathcal{L}$  [14, p. 19].

**Definition 2:** A matching  $\mu$  is deemed to be blocked by an individual  $PU_l$  (or  $CU_k$ ), where we have  $l \in \mathcal{L}$  (or  $k \in \mathcal{K}$ ).

TABLE III  
THE ALGORITHM FOR CASE 3 SCENARIO

---

**Step 1: Initialization**

- 1) Set  $\alpha_{l,k} = \alpha_{l,k}^{cen}$  of Eq. (17), and set  $\tau$ .
- 2) Compute  $C_{CU}^*$  of Eq. (17).
- 3) Compute  $C_{PU}^*$  of Eq. (17).

**Step 2: Select the players to act as the proposer.**

- Compare the value of  $C_{CU}^*$  and  $C_{PU}^*$ .
- if  $C_{CU}^* < C_{PU}^*$ , Select the PUs to act as the proposers and use the DMA in Case 1 of Table I.
- Else, select the CUs to act as the proposers and use the DMA in Case 2 of Table II.
- Set  $l = l + 1, k = k + 1$ .

**until converge**

**Step 3: End of Algorithm**

---

The  $PU_l$  (or  $CU_k$ ) would rather prefer not to be matched, instead of being matched to its current partner candidate under  $\mu$  [14, p. 20], when a low resultant rate for  $PU_l$  (or  $CU_k$ ) under the current matching  $\mu$  fails to satisfy their minimum rate requirements at a given power allocation factor  $\alpha_{l,k}$ . This mathematically implies that:

$$\begin{aligned} R_{l,\mu(l)}^{PU}(\alpha_{l,\mu(l)}) &< R_{l,req}^{PU} \quad \text{or} \\ R_{\mu(k),k}^{CU}(\alpha_{\mu(k),k}) &< R_{k,req}^{CU}. \end{aligned} \quad (25)$$

**Definition 3:** The pair  $(PU_l, CU_k)$  is referred to as a blocking pair for the matching  $\mu$ , if both  $PU_l$  and  $CU_k$  would prefer to be matched with each other, but they are not matched under the current matching  $\mu$  [14, p. 21]. Both  $PU_l$  and  $CU_k$  achieve a higher rate if they are matched in comparison with their current matching partners under  $\mu$  at a given power-allocation factor  $\alpha_{l,k}$ . This mathematically implies that:

$$\begin{aligned} R_{l,k}^{PU}(\alpha_{l,k}) &> R_{l,\mu(l)}^{PU}(\alpha_{l,\mu(l)}) \quad \text{and} \\ R_{l,k}^{CU}(1 - \alpha_{l,k}) &> R_{\mu(k),k}^{CU}(1 - \alpha_{\mu(k),k}). \end{aligned} \quad (26)$$

**Definition 4:** A matching  $\mu$  is stable if it is not blocked by any individual or any pair [14, p. 21].

Based on Definition 1, 2, 3 and 4, we now proceed to prove that our DMA as shown in Section IV-B constitutes a stable matching.

Let us assume that the final matching result provided by our DMA is  $\mu$ . There are no blocking individuals under  $\mu$  because all users in each preference list (Eq.(20) and Eq.(22)) satisfy the corresponding minimum rate requirements. Furthermore, we assume that  $PU_l$  and  $CU_q$  ( $q \in K$ ) are not matched with each other under  $\mu$ , but  $PU_l$  prefers to be matched with  $CU_q$  instead of its current partner  $CU_{\mu(l)}$  under  $\mu$ . This illustrates that  $CU_q$  is acceptable to  $PU_l$ , hence  $PU_l$  must have made an offer to  $CU_q$  before asking its current partner  $CU_{\mu(l)}$  (or before being rejected by all its acceptable partners). Since  $PU_l$  was not matched to  $CU_q$ , when the algorithm was terminated, this implies that  $PU_l$  must have been rejected by  $CU_q$ , where  $CU_q$  is in favor of  $PU_t$  ( $t \neq l, t \in L$ ), which offers a better rate compared to that offered by  $PU_l$ , as detailed in Step-2B of Table I. Therefore,  $CU_q$  is matched to its current partner  $PU_t$ , where  $t = \mu(q)$ , which is better than  $PU_l$ . As a result,

neither  $PU_l$  nor  $CU_q$  blocks the matching  $\mu$ , hence the DMA is stable.

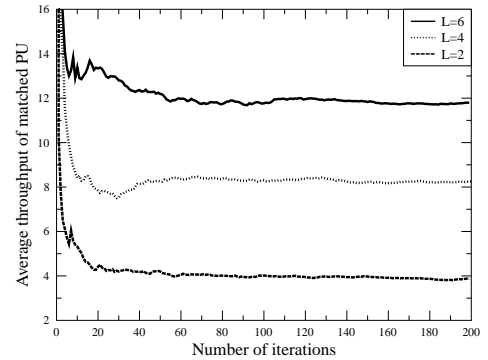


Fig. 3. The performance of average throughput of matched PUs versus the number of iterations by employing the conventional DMA in the CR-NOMA systems.

Fig. 3 shows the average throughput of matched PUs versus the number of iterations  $N$  for the DMA in CR-NOMA system. It implies that the DMA converges to a certain value after a finite number of iterations. From Fig. 3, it can be noted that the DMA converges to the identical throughput of  $R_{match}^{PU}$  at  $L = 2, L = 4$  and  $L = 6$ . Furthermore, we can observe that the number of iterations required for converge slightly increases as  $L$  gets bigger. That is because, the DMA would use more iterations to process the matching procedure between the PUs and CUs, while we have more number of user pairs. Therefore, we can imply that, our DMA requires low implementation complexity, which will be further analyzed in Section V-B.

### B. Complexity Analysis

The complexity required by the CA and the conventional DMA will be discussed in this section in terms of the number of operations [24]. The optimal CA requires the exhaustive search. All possible matching pairs over PUs and CUs are searched, and then the most profitable matched pair is selected according to the objective function as discussed in Eq. (8) of Section III. The complexity of the CA can be expressed as:

$$\mathcal{O} ( L! 2^{L+K} ), \quad (27)$$

where the  $L!$  of Eq. (27) is the total number of matching combinations between the PUs and CUs for the case of  $L = K$ . Additionally, the complexity of solving the linear programming problem for all possible matching combinations is given by:  $2^{L+K}$  [25].

For the conventional DMA, the PUs and CUs negotiate with each other for transmitting power as discussed in Section IV-B. In the DMA, the PUs make the offers to the CUs, which then decide either to accept or reject the offers. The matched CU rejects its current matching in favor of a new matching, whenever the new matching pair is capable of providing a higher rate for this CU. In that case, the rejected PU updates its power allocation factor according to  $\alpha'_{l,cur,k} = \alpha_{l,cur,k} - \tau$ . The worst case happens for a PU is that, it updates its preference list ( $\frac{\alpha_{init} - \alpha_{limit}}{\tau}$ ) times until it finds a final matched partner. The  $\alpha_{limit}$  is the tolerable minimum power allocation factor



that can be accepted by this PU. Note that, the PU's achievable rate is less than its minimum rate requirement when  $\alpha_{l,k} < \alpha_{limit}$ . Let us consider the most worst case which requires the highest complexity. A PU has  $K$  CUs in its preference list. This PU offers  $(\frac{\alpha_{init} - \alpha_{limit}}{\tau})$  times for each of the  $(K - 1)$  CUs in its preference list. Then, the PU finds a final matched CU when the PU makes the last offer to the CU in the bottom of its list. For this extreme case, the number of operations required should be  $(\frac{\alpha_{init} - \alpha_{limit}}{\tau}) \times K(K-1)$ , which corresponds to the complexity of  $\mathcal{O} = \{K^2\}$ . Therefore, the complexity of the DMA is given by

$$\mathcal{O} ( LK^2 ) . \quad (28)$$

It can be seen from Eq. (27) and Eq. (28) that, the CA has a significantly higher complexity compared with the DMA, since the complexity of CA increases exponentially by the number of PUs or CUs.

### C. Rate Analysis

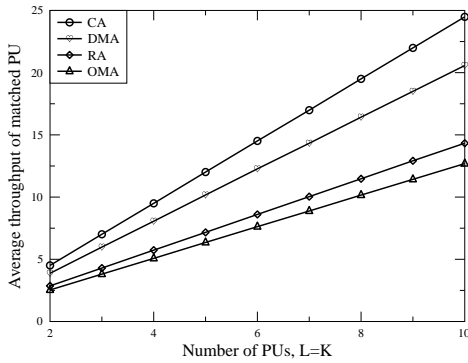


Fig. 4. The performance of the average throughput of matched PUs versus the number of PUs in the CR-NOMA system over Rayleigh fading channel. The objective function of case 1 is considered, which is described in Eq. (9) of Section III. The ‘‘CA’’, ‘‘DMA’’ and ‘‘RA’’ techniques are detailed in Section IV-B. The transmit SNRs of the CUs are  $\gamma^{CU} = 5$  dB and the transmit SNRs of the PUs are  $\gamma^{PU} = 3$  dB.

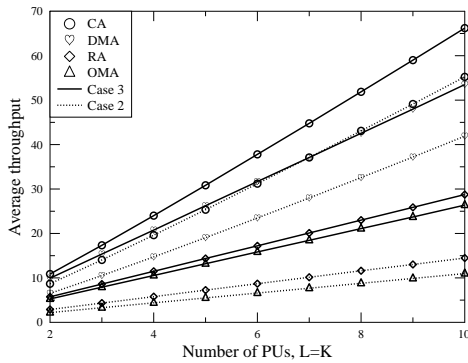


Fig. 5. Performance of the average throughput versus the number of PUs in the CR-NOMA system over Rayleigh fading channel. The objective functions of case 2 and case 3 are considered, which are described in Eq. (10) and Eq. (11) of Section III.

In our current design work, we have considered multiple pairs of users, which the maximum number of pairs is  $M = 10$ . Note that, both our proposed algorithms are scalable and

can be used in large CR-NOMA systems. In our simulations, we assume that the transmit SNRs of all CUs are equal, and set  $\gamma_1^{CU} = \dots = \gamma_k^{CU} = 5$  dB. We also assume that the SNRs of all PUs are  $\gamma_1^{PU} = \dots = \gamma_l^{PU} = 3$  dB. The rate requirement of the CUs are assumed to  $R_{k,req}^{CU} = 2$ . Fig. 4 shows the average sum-rate of the matched PUs versus the number of PUs in the CR-NOMA system, when  $L = K = \{2, 3, \dots, 10\}$ . The average sum-rate of the matched PUs is defined as:

$$R_{match}^{PU} = \sum_l \sum_k m_{l,k_{matched}} \times R_l^{PU}(\alpha_{l,k_{matched}}) , \quad (29)$$

where  $\alpha_{l,k_{matched}}$  represents the final value of power allocated to the matched PU. In our simulations, we evaluate the bandwidth normalized throughput of the systems. In Fig. 4 we compare the performances achieved by the three algorithms, which include the DMA of Section IV-B, the CA described in Section III-B and another benchmark, namely random algorithm (RA). In the context of the RA, each PU makes an offer  $\alpha_{l,k}$  to a CU, which is randomly selected from its preference list. The CU chooses the specific matching pair that provides an increased rate to itself and discards the one having a lower rate. More specifically, the  $\alpha_{l,k}$  is chosen in order to maximize the optimization weight. In Fig. 4, we have considered the scenario for case 1 of Eq. (9) as introduced in Section III. In our simulations, we used the optimal power allocation factor for the RA according to different objective functions as shown in Eq. (9), Eq. (10) and Eq. (11).

In Fig. 4, we observe that the CA achieves the highest average sum-rate among these three algorithms, while the RA achieves the lowest sum-rate. Furthermore, as seen from Fig. 4 that our CR-NOMA system always achieve a higher sum-rate than that of the OMA system, regardless of the user pair and power allocation approaches used. Specifically, when  $L = 6$ , for the DMA under the CR-NOMA system, the average throughput of matched PUs is  $12.5 - 7.5 = 5$  bit-per-second-per-Hz (bit/s/Hz), which is higher than that of the OMA scheme. Moreover, we find that our CR-NOMA system achieves a higher performance gain over the OMA system as the value of  $L$  increases. Even when the RA is employed under the CR-NOMA, it still outperforms the OMA case, such as  $9.2 - 7.5 = 1.7$  bit/s/Hz can be obtained at  $L = 6$ .

Additionally, Fig. 5 shows the average throughput of matched PUs and CUs as well as the sum-rate of matched CUs versus the number of PUs in the CR-NOMA system. In Fig. 5, we consider the performances for the two optimization objective functions, which are the case 2 of Eq. (9) and case 3 of Eq. (11) as described in Section III. The average sum-rate of matched CUs is formulated as:

$$R_{match}^{CU} = \sum_l \sum_k m_{l,k_{matched}} \times R_k^{CU}(1 - \alpha_{l,k_{matched}}) . \quad (30)$$

Then the average sum-rate of matched PUs and CUs is given by:

$$R_{match}^{total} = \sum_l \sum_k m_{l,k_{matched}} \times [R_k^{PU}(\alpha_{l,k_{matched}}) + R_k^{CU}(1 - \alpha_{l,k_{matched}})] . \quad (31)$$

As shown in Fig. 5, for both two optimization objective functions, maximizing the average throughput of matched PUs

and CUs, and maximize the throughput of matched CUs, the CA achieves the best performance. We observe from Fig. 5 that, the performance of our DMA is close to that of the CA, which implies the DMA provides a promising suboptimal solution to the problem in Eq. (8). However, the RA achieves the worst performance among the three algorithms in the CR-NOMA system. Once again, Fig. 5 illustrates our CR-NOMA system can outperform the OMA system under all cases considered.

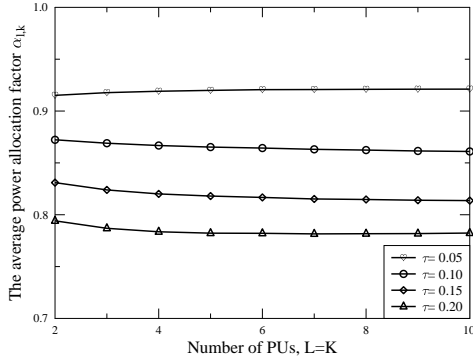


Fig. 6. The performance of the average power allocation factor  $\alpha_{l,k}$  versus the step size of the power allocation factor  $\tau$  in the CR-NOMA system.

In Fig. 6, we investigate how the step size  $\tau$  affects the power allocation, i.e.  $\alpha_{l,k}$ . As expected, the power allocated factor  $\alpha_{l,k}$  was decreasing with increasing  $\tau$ . However, a small  $\tau$  would lead to a longer matching period. Hence, we have chosen  $\tau = 0.05$  as a compromise in our simulations.

#### D. Relative Frequency of Successfully matched PUs

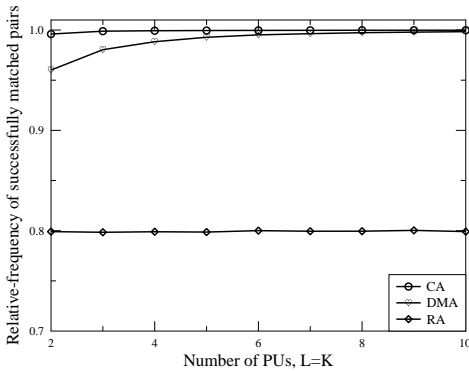


Fig. 7.  $P_{match}$  versus  $L$  performance of the CR-NOMA system over Rayleigh fading channel. The “CA”, “DMA” and “RA” techniques are detailed in Section IV-B. The transmit SNRs of the CUs are  $\gamma^{CU} = 5$  dB and the transmit SNRs of the PUs are  $\gamma^{PU} = 3$  dB.

Fig. 7 illustrates the relative frequency of successfully matched PUs, i.e.  $P_{match}$ , versus the number of PUs, and  $P_{match}$  is evaluated as:

$$P_{match} = \frac{1}{L} E \left[ \sum_l^L \sum_k^K m_{l,k} \right], \quad (32)$$

where  $E[\cdot]$  denotes the expected value of  $[\cdot]$ . As shown in Fig. 7,  $P_{match}$  of the CA is the highest, followed by that of

the DMA, and that of the RA is the worst. Hence, the rate-improvement observed in Fig. 4 is linked to a higher  $P_{match}$ , resulting from a better matching. Therefore,  $P_{match}$  of the DMA is identical to that of the CA when  $L = K = 5$ . Additionally,  $P_{match}$  of the RA achieves an average value, which is  $P_{match} = 0.8$  for all different numbers of the PUs and CUs. Moreover, the percentage of matched users found by the DMA is seen to be significantly higher than that of the RA. For example, in Fig. 7, the DMA achieves  $P_{match} = 99.99\%$  at  $K = 10$ , while the RA only achieves  $P_{match} = 80\%$ . Thus there is an approximately  $P_{match} = 99.99\% - 80\% = 19.99\%$  difference between these two algorithms.

#### E. Imperfect channel estimation

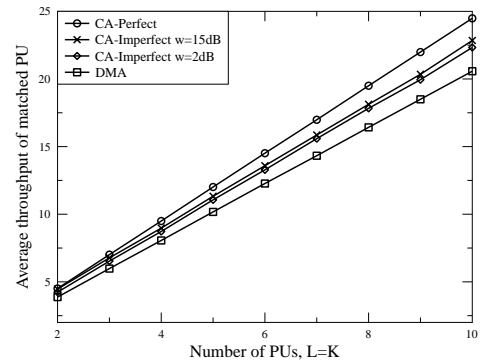


Fig. 8. Performance of the average throughput versus the number of PUs of CR-NOMA system by implementing the imperfect channel estimations, associate with the error estimation factor  $w = 2$ dB and  $w = 15$ dB.

However perfect CSI estimation is not realistic in practice, especially in a mobile communication system having time-varying wireless channels. In this section, we will investigate the performance of the considered NOMA-CR network, when the CSI estimation is impaired by estimation errors. We employ the widely used assumptions [26], [27] in which the feedback to the transmitter is instantaneous and error free. More specifically, the CSI is also achievable at the transmitter whatever CSI the receiver has. In the case of imperfect channel knowledge, the estimate of the channel gain matrix  $\tilde{H}$  is constructed by the channel gain matrix  $H$  modeled by independent zero-mean complex-valued Gaussian random variables and the channel error matrix  $\tilde{N}$  that has a variance of  $var(\tilde{N})$ . The estimation of the channel gain matrix is written as [28]:

$$\tilde{H} = H + \tilde{N}. \quad (33)$$

The channel  $H$  is defined in Eq. (1) and Eq. (2). The channel estimation error  $\tilde{N}$  is assumed to be independent identical Gaussian distributed obeying the standard Gaussian distribution of zero mean. The degree of CSI estimation errors is governed by the Channel Estimation Factor (CEF), namely  $w$  (dB) defined by:

$$w = 10 \log_{10} \frac{1}{var(\tilde{N})} [\text{dB}]. \quad (34)$$

Then the received signal of our considered NOMA-CR system impaired by imperfect channel estimation is rewritten as:

$$\begin{aligned} Y &= \tilde{H}X + n \\ &= (H + \tilde{N})X + n. \end{aligned} \quad (35)$$

However, we emphasize that, regardless of the estimation techniques employed, the estimation error is actually non-Gaussian, hence using a Gaussian model constitutes an approximation. The simulation results of our proposed scheme subjected to CSI estimation channel errors are discussed in the following. Fig. 8 shows a comparison between the perfect channel estimation for the average throughput and the imperfect channel estimation results associated with the CSI estimation errors of  $w = 2\text{dB}$  and  $w = 15\text{dB}$ . It is evident that the average throughput under the perfect CSI is higher than that of imperfect CSI scenarios, when the other parameters stay the same. Furthermore, the average throughput of PU achieved by the CA gets worse under the imperfect CSI, when the value of the error estimation factor decreases, i.e. when error estimation becomes more severe. It is worth noting that, when considering imperfect CSI, the throughput performance of our proposed DMA gets much closer to that of the CA, which reveals that our distributed algorithm, i.e. the DMA is a promising solution to practical implementation of user pairing and power allocation in practical CR-NOMA systems.

## VI. PRACTICAL ATTCM AIDED CR-NOMA NETWORK

### A. ATTCM scheme

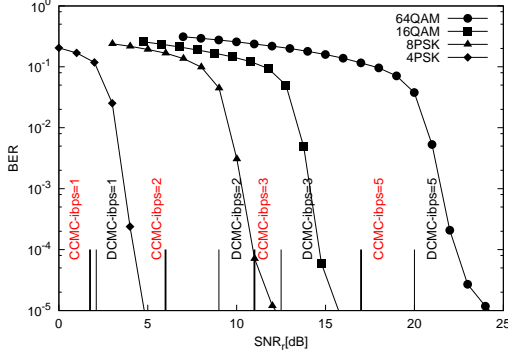


Fig. 9. The BER versus  $\text{SNR}_r$  performance of TTCM using a frame length of 120,000 symbols, when communicating over Rayleigh fading channels. Four TTCM iterations were invoked.

In this section we investigate the practical design of the CR-NOMA system advocated using ATTCM. Employing TTCM has the advantage that the system's effective throughput can be increased upon increasing the code rate, when the channel-quality improves. Additionally, the Bit error ratio (BER) performance of the system may be improved when TTCM is used [29]. The TTCM encoder comprises two identical parallel-concatenated TCM encoders [30] linked by a symbol interleaver. The first TCM encoder directly processes the original input bit sequence, while the second TCM encoder manipulates the interleaved version of the input bit sequence. Then the bit-to-symbol mapper maps the input bits to complex-valued ATTCM symbols using the Set Partition based labelling method [29]. The structure of the TTCM decoder is similar

to that of binary turbo codes, but each decoder alternately processes its corresponding encoder's channel-impaired output symbol, and then the other encoder's channel-impaired output symbol [29, p.764]. More details on the ATTCM principles can be found in [29].

We have employed a ATTCM scheme for the CR-NOMA system, where the information Bit-per-Symbol (iBPS) is selected from the set  $\text{iBPS} = \{0, 1, 2, 3, 5\}$  Bit-per-Symbol (BPS) corresponding to the case of no transmission, QPSK, 8PSK, 16QAM and 64QAM modulations, respectively. Moreover, the ATTCM mode switching thresholds  $\Upsilon = [\gamma_0, \gamma_1, \gamma_2, \gamma_3]$  were determined based on the BER performance curves of each of the four TTCM schemes in a Rayleigh fading channel, which is shown in Fig. 9. Specifically, the ATTCM mode switching operation and the throughput of the modes are specified by the following algorithm:

$$\text{MODE} = \begin{cases} \gamma_3 \leq \gamma_R, & \text{TTCM-64QAM, iBPS=5 BPS;} \\ \gamma_2 \leq \gamma_R < \gamma_3, & \text{TTCM-16QAM, iBPS=3 BPS;} \\ \gamma_1 \leq \gamma_R < \gamma_2, & \text{TTCM-8PSK, iBPS=2 BPS;} \\ \gamma_0 \leq \gamma_R < \gamma_1, & \text{TTCM-4PSK, iBPS=1 BPS;} \\ \gamma_R < \gamma_0, & \text{No transmission, iBPS=0 BPS.} \end{cases}$$

We use the notation  $\gamma_R^{PU}$  to refer to the instantaneous received SNR of the link between the BS and the  $\text{PU}_l$  and  $\gamma_R^{CU}$  represents the instantaneous received SNR of the link between the BS and the  $\text{CU}_k$ . The received SNR at  $\text{PU}_l$  is given by

$$\gamma_R^{PU} = \frac{\alpha_{l,k} P_S \varrho_p |h_{B,p_l}|^2}{(1 - \alpha_{l,k}) P_S + N_0}. \quad (36)$$

The received SNR at  $\text{CU}_k$  can be expressed as:

$$\gamma_R^{CU} = \frac{(1 - \alpha_{l,k}) P_S \varrho_c |h_{B,c_k}|^2}{N_0}, \quad (37)$$

where the term  $\alpha_{l,k}$  is the power allocation coefficient of  $\text{PU}_l$ ,  $P_S$  is the transmit power emit from the BS and the pathloss  $\varrho$  is detailed in Section II. The quasi-static Rayleigh fading channels between the BS and  $\text{PU}_l$  is denoted as  $h_{B,p_l}$ , while this between the BS and  $\text{CU}_k$  is represented by  $h_{B,c_k}$ . The channel gains are independent of each other.

Each of the communication links in our CR-NOMA system will be assisted by the ATTCM scheme. We chose the switching thresholds to ensure that the target BER is lower than  $10^{-5}$ , which is given by  $\Upsilon_{\text{ATTCM}} = [4.8, 12, 16, 24]$  dB as seen in Fig. 9. In contrast to the practical power- and bandwidth-efficient ATTCM scheme, we also make use of both continuous-input continuous-output memoryless channel (CCMC) and the discrete-input continuous-output memoryless channel (DCMC) [31] schemes. More specifically, the CCMC based adaptive scheme assumes that idealistic coding and modulation schemes are employed for communicating exactly at Shannon's capacity. By contrast, the DCMC based adaptive scheme assumes that an idealistic capacity-achieving code is employed for allowing the PSK/QAM modulation schemes considered to operate right at the modulation-dependent DCMC capacity. We note that Shannon's CCMC capacity is only restricted by the SNR and the bandwidth. The CCMC-based switching thresholds are represented as  $\Upsilon_{\text{CCMC}} = [1.75, 6, 11, 14]$  dB, while the switching thresholds

of the corresponding modulation dependent DCMC based scheme are given by  $\Upsilon_{DCMC}=[2, 8, 12.5, 20]$  dB, which are also explicitly shown in Fig. 9. The reason why we have chosen the target BER to be lower than  $10^{-5}$  is because the error floor emerging at  $BER < 10^{-5}$  can be removed by using a long outer code, such as a Reed Solomon (RS) code, albeit no RS code was used here. For quasi-static fading channels, the achievable rates over different links become random and vary as the channel changes. The relatively performances of our ATTCM aided CR-NOMA network will be discussed in Section VI-B.

### B. Performance of our ATTCM aided CR-NOMA Network

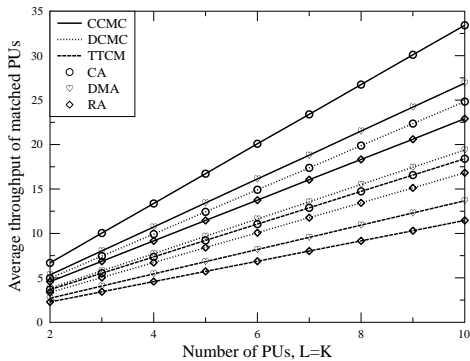


Fig. 10. The performance of  $R_{matched}^{PU}$  versus the number of PUs for the ATTCM, CCMC and DCMC aided CR-NOMA systems over block fading Rayleigh channel. A BER below  $10^{-5}$  is maintained. The “CA”, “DMA” and “RA” techniques are detailed in Section IV-B. The transmit SNRs of the PUs are  $\gamma^{PU} = 10$  and the transmit SNRs of the CUs are  $\gamma^{CU} = 15$  dB.

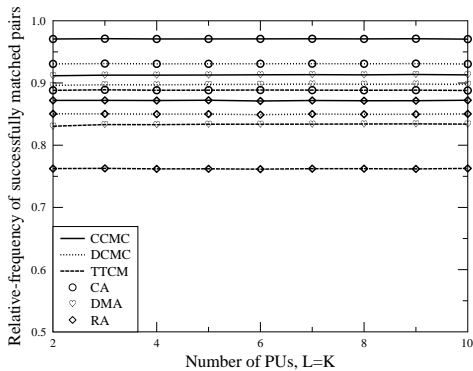


Fig. 11.  $P_{match}$  versus  $L$  performance of the ATTCM, CCMC and DCMC aided CR-NOMA systems over block fading Rayleigh channel. A BER below  $10^{-5}$  is maintained. The “CA”, “DMA” and “RA” techniques are detailed in Section IV-B. The transmit SNRs of the PUs are  $\gamma^{PU} = 10$  and the transmit SNRs of the CUs are  $\gamma^{CU} = 15$  dB.

In this section, we investigate the performance of the CCMC, DCMC and ATTCM aided CR-NOMA systems by employing the DMA, aiming to find the appropriate power allocations of the PU and CU in the same cluster. In Fig. 10, we show that the rate performance of the CCMC, DCMC and ATTCM aided CR-NOMA systems in terms of the average sum-rate of matched PUs. As seen from Fig. 10, when  $L$  increases, the sum-rate of matched PU pairs also increases. That is because the number of matching pairs gets bigger as the total number of PUs increases. Thus, the probability

of successful matching of the PUs and CUs becomes higher, which results in a higher sum-rate. As expected, the DMA algorithm using CCMC achieves a higher average throughput than that for employing the DCMC, as shown in Fig. 10. When we consider the ATTCM scheme, there is a minor difference of  $(7.3 \text{ BPS} - 5.3 \text{ BPS}) = 2 \text{ BPS}$  between the DMA and the CA. Additionally, the performance by employing the DMA in CR-NOMA is much better than the RA. The percentage  $P_{match}$  versus the number of PUs,  $L$ , is investigated in Fig. 11. The DMA based on CR-NOMA achieve higher percentage compare with the RA when they have the same number of PUs. Moreover, the CR-NOMA employing the DMA imposes a lower complexity compared with the system using the CA, and its performance can approach to that of the CA.

## VII. CONCLUSIONS

In this paper, we have applied matching theory to solve the user pairing and power allocation problem in the CR-NOMA systems. In our proposed distributed matching algorithm, the PUs trade the available power with the CUs by negotiating the power allocation coefficients, which guarantees that the rate requirements of both the PUs and CUs are satisfied. We have shown that the proposed DMA results in a stable matching and implement a low complexity. Our numerical analysis has revealed that the DMA achieved a better performance than the RA benchmark scheme, and it is close to the optimum CA. Importantly, we have also shown that the CR-NOMA system can achieve a significant performance advantage over the OMA system. Moreover, we have investigated a practical ATTCM for the CR-NOMA system, and the system’s throughput can be further improved according to our simulation results.

## APPENDIX A

As discussed in Section III, based on Eq. (8) and Eq. (11), the objective function that maximizes the total sum-rate of matched PU and CU can be rewritten as:

$$\max_{\alpha_{l,k}} W(\alpha_{l,k}) = \max_{\alpha_{l,k}} \left[ R_l^{PU}(\alpha_{l,k}) + R_k^{CU}(1 - \alpha_{l,k}) \right]. \quad (38)$$

The term  $R_l^{PU}(\alpha_{l,k})$  is defined in Eq. (3) and  $R_k^{CU}(1 - \alpha_{l,k})$  is expressed in Eq. (4). The problem of Eq. (38) is a concave function problem and its relative proof is discussed in Appendix B. Additionally, its three linear constraints which are the condition (a), (b) and (c) of Eq. (8) are shown as follows:

$$s.t. \quad (a) \quad R_l^{PU}(\alpha_{l,k}) \geq R_{l,req}^{PU}, \forall l \forall k \quad (39)$$

$$(b) \quad R_k^{CU}(\alpha_{l,k}) \geq R_{k,req}^{CU}, \forall l \forall k \quad (40)$$

$$(c) \quad 0 \leq \alpha_{l,k} \leq 1, \forall l \forall k. \quad (41)$$

Hence, the Lagrangian function for Eq. (38) using the corresponding constrains of Eq. (39), Eq. (40) and Eq. (41) can be expressed as:

$$\begin{aligned} \mathcal{L}(\alpha_{l,k}, \lambda, \mu) = & W(\alpha_{l,k}) \\ & + \lambda_{l,k} \left( R_l^{PU}(\alpha_{l,k}) - R_{l,req}^{PU} \right) \\ & + \mu_{l,k} \left( R_k^{CU}(1 - \alpha_{l,k}) - R_{k,req}^{CU} \right), \end{aligned} \quad (42)$$

where the weight  $W(\alpha_{l,k})$  in Eq. (38) aims to maximize the sum-rate of the PU and CU.  $\lambda = (\lambda_{l,k} : l \in L, k \in K)$  is a matrix of Lagrange multipliers corresponding to the PUs' rate requirement constraint of Eq. (39) with  $\lambda_{l,k} \geq 0$ . Then the vector  $\mu = (\mu_{l,k} : l \in L, k \in K)$  of Lagrange multipliers corresponds to the CUs' rate constraint of Eq. (40) with  $\mu_{l,k} \geq 0$ . The maximization problem of Eq. (42) can be simplified as:

$$\begin{aligned} \mathcal{L}(\alpha_{l,k}^{cen}, \lambda, \mu) = & \left( \log_2 \left( \frac{\gamma_l^{PU} |h_l^p|^2 + 1}{\gamma_l^{PU} |h_l^p|^2 - \alpha_{l,k} \gamma_l^{PU} |h_l^p|^2 + 1} \right) \right. \\ & + \log_2 (1 + \gamma_k^{CU} |h_k^c|^2 - \alpha_{l,k} \gamma_k^{CU} |h_k^c|^2) \\ & + \lambda_{l,k} \left( \log_2 \left( \frac{\gamma_l^{PU} |h_l^p|^2 + 1}{\gamma_l^{PU} |h_l^p|^2 - \alpha_{l,k} \gamma_l^{PU} |h_l^p|^2 + 1} \right) - R_{l,req}^{PU} \right) \\ & \left. + \mu_{l,k} \left( \log_2 (1 + \gamma_k^{CU} |h_k^c|^2 - \alpha_{l,k} \gamma_k^{CU} |h_k^c|^2) - R_{k,req}^{CU} \right) \right). \end{aligned} \quad (43)$$

The optimization allocation of  $\alpha_{l,k}^{cen}$  for fixed values of  $\lambda_{l,k}$  and  $\mu_{l,k}$  can be calculated for each PU and CU by applying the Karush-Luhn-Tucker (KKT) conditions [32], such that we have:

$$\begin{aligned} \frac{\partial \mathcal{H}(\alpha_{l,k}^{cen}, \lambda, \mu)}{\partial \alpha_{l,k}} = & \frac{1 + \lambda_{l,k}}{\gamma_l^{PU} |h_l^p|^2 - \alpha_{l,k} \gamma_l^{PU} |h_l^p|^2 + 1} \\ & - \frac{1 - \mu_{l,k}}{1 + \gamma_k^{CU} |h_k^c|^2 - \alpha_{l,k} \gamma_k^{CU} |h_k^c|^2}. \end{aligned} \quad (44)$$

Then the Lagrange multipliers corresponding to the constraint of Eq. (39) is given by

$$\begin{aligned} \frac{\partial \mathcal{H}(\alpha_{l,k}^{cen}, \lambda, \mu)}{\partial \lambda_{l,k}} = & \log_2 \left( \frac{\gamma_l^{PU} |h_l^p|^2 + 1}{\gamma_l^{PU} |h_l^p|^2 - \alpha_{l,k} \gamma_l^{PU} |h_l^p|^2 + 1} \right) \\ & - R_{l,req}^{PU}, \end{aligned} \quad (45)$$

where  $R_{l,req}^{PU}$  is defined in Eq. (5). Then the optimization value of  $\alpha_{l,k}$  based on Eq. (45) can be obtained as:  $\alpha_{l,k} = \frac{\gamma_l^{PU} |h_l^p|^2 + 1 - \sqrt{\gamma_l^{PU} |h_l^p|^2 + 1}}{\gamma_l^{PU} |h_l^p|^2}$ . Additionally, the Lagrange multipliers  $\mu_{l,k}$  corresponds to the constraint of Eq. (40) can be expressed as:

$$\begin{aligned} \frac{\partial \mathcal{H}(\alpha_{l,k}^{cen}, \lambda, \mu)}{\partial \mu_{l,k}} = & \log_2 (1 + \gamma_k^{CU} |h_k^c|^2 - \alpha_{l,k} \gamma_k^{CU} |h_k^c|^2) - R_{k,req}^{CU} \\ \rightarrow \alpha_{l,k} = & \frac{1 + \gamma_k^{CU} |h_k^c|^2 - 2R_{k,req}^{CU}}{\gamma_k^{CU} |h_k^c|^2}. \end{aligned} \quad (46)$$

Hence, we could obtain two bounds on the optimization value for  $\alpha_{l,k}^{cen}$ . Based on Eq. (44),  $\lambda$  and  $\mu$  can be expressed as:

$$\lambda_{l,k} = (1 + \mu_{l,k}) \frac{\gamma_l^{PU} |h_l^p|^2 - \alpha_{l,k} \gamma_l^{PU} |h_l^p|^2 + 1}{1 + \gamma_k^{CU} |h_k^c|^2 - \alpha_{l,k} \gamma_k^{CU} |h_k^c|^2} - 1, \quad (47)$$

$$\mu_{l,k} = (1 + \lambda_{l,k}) \frac{\gamma_k^{CU} |h_k^c|^2 - \alpha_{l,k} \gamma_k^{CU} |h_k^c|^2 + 1}{1 + \gamma_l^{PU} |h_l^p|^2 - \alpha_{l,k} \gamma_l^{PU} |h_l^p|^2} - 1. \quad (48)$$

Since  $\lambda_{l,k} > 0$  and  $\mu_{l,k} > 0$ , Eq. (47) and Eq. (48) may be further derived as:

$$\mu_{l,k} > \frac{\overbrace{C_{CU}^*}}{\overbrace{C_{PU}^*}} \frac{\gamma_{CU} |h_k^c|^2 - \alpha_{l,k} \gamma_{CU} |h_k^c|^2 + 1}{1 + \gamma_{PU} |h_l^p|^2 - \alpha_{l,k} \gamma_{PU} |h_l^p|^2} - 1, \quad (49)$$

$$\lambda_{l,k} > \frac{\overbrace{C_{PU}^*}}{\overbrace{C_{CU}^*}} \frac{1 + \gamma_{PU} |h_l^p|^2 - \alpha_{l,k} \gamma_{PU} |h_l^p|^2}{\gamma_{CU} |h_k^c|^2 - \alpha_{l,k} \gamma_{CU} |h_k^c|^2 + 1} - 1. \quad (50)$$

Moreover, when the conditions in Eq. (49) and Eq. (50) are satisfied, we could obtain the bounds related to Eq. (45) and Eq. (46). Finally, the optimization value of  $\alpha_{l,k}^{cen}$  that maximizes the total sum-rate of PU and CU is given by

$$\alpha_{l,k}^{cen} = \begin{cases} \frac{1 + \gamma_k^{CU} |h_k^c|^2 - 2R_{k,req}^{CU}}{\gamma_l^{PU} |h_l^p|^2 + 1 - \sqrt{\gamma_l^{PU} |h_l^p|^2 + 1}} & C_{CU}^* < C_{PU}^*; \\ \frac{\gamma_l^{PU} |h_l^p|^2 + 1 - \sqrt{\gamma_l^{PU} |h_l^p|^2 + 1}}{\gamma_l^{PU} |h_l^p|^2} & C_{CU}^* > C_{PU}^*; \end{cases} \quad (51)$$

## APPENDIX B

According to Eq. (3) and Eq. (4), the optimization problem of Eq. (38) can be rewritten as:

$$\begin{aligned} \max_{\alpha_{l,k}} \log_2 & \left( 1 + \frac{\alpha_{l,k} \gamma_l^{PU} |h_l^p|^2}{(1 - \alpha_{l,k}) \gamma_l^{PU} |h_l^p|^2 + 1} \right) \\ & + \log_2 \left( 1 + \gamma_k^{CU} |h_k^c|^2 - \alpha_{l,k} \gamma_k^{CU} |h_k^c|^2 \right). \end{aligned} \quad (52)$$

$f(X_P)$   $f(X_C)$

The second-derivation of Eq. (52) is derived as:

$$\begin{aligned} [f(X_P) + f(X_C)]'' = & \frac{\overbrace{\left( \frac{(\gamma_{PU} |h_l^p|^2)^2}{\ln 2} [\gamma_l^{PU} |h_l^p|^2 + 1 - \alpha_{l,k} \gamma_l^{PU} |h_l^p|^2]^{-2} \right)}^{f(X_A)}}{\overbrace{\left( \frac{(\gamma_k^{CU} |h_k^c|^2)^2}{\ln 2} [\gamma_k^{CU} |h_k^c|^2 + 1 - \alpha_{l,k} \gamma_k^{CU} |h_k^c|^2]^{-2} \right)}^{f(X_B)}} \end{aligned} \quad (53)$$

As shown in Eq. (53), the term  $f(X_A) > 0$  and the term  $f(X_B) < 0$ . In our system, we assume that the term  $f(X_B)$  is always greater than  $f(X_A)$ , thus  $f(X_A) + f(X_B) < 0$ . Therefore, the function  $[f(X_P) + f(X_C)]'' < 0$ , which shows that the problem of Eq. (52) is a concave problem [33].

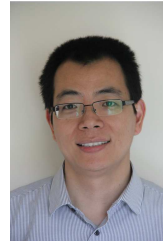
## REFERENCES

- [1] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," in *IEEE 77th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, June 2013.
- [2] S. L. Shieh and Y. C. Huang, "A Simple Scheme for Realizing the Promised Gains of Downlink Nonorthogonal Multiple Access," *IEEE Transactions on Communications*, vol. 64, pp. 1624–1635, April 2016.
- [3] J. Choi, "Minimum Power Multicast Beamforming With Superposition Coding for Multiresolution Broadcast and Application to NOMA Systems," *IEEE Transactions on Communications*, vol. 63, pp. 791–800, March 2015.
- [4] Z. Ding, M. Peng, and H. Poor, "Cooperative Non-Orthogonal Multiple Access in 5G Systems," *IEEE Communications Letters*, vol. 19, pp. 1462–1465, Aug 2015.
- [5] Z. Ding, F. Adachi, and H. Poor, "The Application of MIMO to Non-Orthogonal Multiple Access," *IEEE Transactions on Wireless Communications*, vol. PP, no. 99, pp. 1–1, 2015.
- [6] H. Marshoud, V. M. Kapinas, G. K. Karagiannidis, and S. Muhaidat, "Non-orthogonal multiple access for visible light communications," *IEEE Photonics Technology Letters*, vol. 28, pp. 51–54, Jan 2016.
- [7] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. PP, no. 99, pp. 1–1, 2016.

- [8] Y. Zhang, H. M. Wang, Q. Yang, and Z. Ding, "Secrecy sum rate maximization in non-orthogonal multiple access," *IEEE Communications Letters*, vol. PP, no. 99, pp. 1–1, 2016.
- [9] Z. Ding, Z. Yang, P. Fan, and H. Poor, "On the Performance of Non-Orthogonal Multiple Access in 5G Systems with Randomly Deployed Users," *IEEE Signal Processing Letters*, vol. 21, pp. 1501–1505, Dec 2014.
- [10] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, pp. 894–914, May 2009.
- [11] Z. Ding, R. Schober, and H. V. Poor, "A General MIMO Framework for NOMA Downlink and Uplink Transmission Based on Signal Alignment," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 4438–4454, June 2016.
- [12] L. Lv, J. Chen, and Q. Ni, "Cooperative Non-Orthogonal Multiple Access in Cognitive Radio," *IEEE Communications Letters*, vol. 20, pp. 2059–2062, Oct 2016.
- [13] D. Gusfield and R. W. Irving, *The Stable Marriage Problem: Structure and Algorithms*. The MIT Press, 1989.
- [14] A. Roth and M. Sotomanyor, *Two sided matching: a study in game-theoretic modeling and analysis, 1st ed.* Cambridge University Press, 1989.
- [15] Y. Gu, W. Saad, M. Bennis, M. Debbah, and Z. Han, "Matching theory for future wireless networks: fundamentals and applications," *IEEE Communications Magazine*, vol. 53, pp. 52–59, May 2015.
- [16] Y. Xiao, K. C. Chen, C. Yuen, Z. Han, and L. DaSilva, "A bayesian overlapping coalition formation game for device-to-device spectrum sharing in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, pp. 4034–4051, July 2015.
- [17] S. Bayat, R. Louie, Z. Han, B. Vucetic, and Y. Li, "Physical-layer security in distributed wireless networks using matching theory," *IEEE Transactions on Information Forensics and Security*, vol. 8, pp. 717–732, May 2013.
- [18] B. Di, S. Bayat, L. Song, and Y. Li, "Radio Resource Allocation for Downlink Non-Orthogonal Multiple Access (NOMA) Networks Using Matching Theory," in *2015 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec 2015.
- [19] K. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. Cambridge University Press, New York, 2009.
- [20] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5g systems," *IEEE Signal Processing Letters*, vol. 22, pp. 1647–1651, Oct 2015.
- [21] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and L. Hanzo, "Enhancing the physical layer security of non-orthogonal multiple access in large-scale networks," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 1656–1672, March 2017.
- [22] B. Zhao and M. C. Valenti, "Distributed turbo coded diversity for relay channel," *Electronics Letters*, vol. 39, pp. 786 – 787, May 2003.
- [23] M. Shamaian, S. H. Lee, S. Vishwanath, and H. Vikalo, "Distributed algorithms for spectrum access in cognitive radio relay networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 1947–1957, November 2012.
- [24] J. Jalden and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Transactions on Signal Processing*, vol. 53, pp. 1474–1484, April 2005.
- [25] A. Kan and J. Telgen, "The complexity of linear programming," *J. Statistica Neerlandica*, no. 2, pp. 91–107, 1981.
- [26] Z. Yang, Z. Ding, P. Fan, and G. K. Karagiannidis, "On the Performance of Non-orthogonal Multiple Access Systems With Partial Channel Information," *IEEE Transactions on Communications*, vol. 64, pp. 654–667, Feb 2016.
- [27] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, pp. 2203–2214, May 2006.
- [28] T. Weber, A. Sklavos, and M. Meurer, "Imperfect channel-state information in MIMO transmission," *IEEE Transactions on Communications*, vol. 54, pp. 543 – 552, March 2006.
- [29] L. Hanzo, S. X. Ng, T. Keller, and W. Webb, *Quadrature Amplitude Modulation: From Basics to Adaptive Trellis-Coded, Turbo-Equalised and Space-Time Coded OFDM, CDMA and MC-CDMA Systems*. Wiley-IEEE Press, 2004.
- [30] G. Ungerböck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, pp. 55–67, January 1982.
- [31] S. Ng and L. Hanzo, "On the MIMO channel capacity of multidimensional signal sets," *IEEE Transactions on Vehicular Technology*, vol. 55, pp. 528 – 536, March 2006.
- [32] H. Hindi, "A tutorial on convex optimization II: duality and interior point methods," in *American Control Conference*, p. 11, June 2006.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

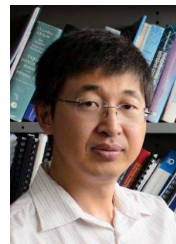


**Wei Liang** received her M.Sc. and Ph.D. degree in wireless communication at University of Southampton, Southampton, U.K in 2010 and 2015, respectively. She is a Research Fellow in Lancaster University now. Her research interests include adaptive coded modulation, network coding, matching theory, game theory, cooperative communication, cognitive radio network and Non-orthogonal multiple access scheme.



**Zhiguo Ding** (S'03-M'05) received his B.Eng in Electrical Engineering from the Beijing University of Posts and Telecommunications in 2000, and the Ph.D degree in Electrical Engineering from Imperial College London in 2005. From Jul. 2005 to Aug. 2014, he was working in Queen's University Belfast, Imperial College and Newcastle University. Since Sept. 2014, he has been with Lancaster University as a Chair Professor. From Oct. 2012 to Sept. 2016, he has also been an academic visitor in Princeton University.

Dr Ding' research interests are 5G networks, game theory, cooperative and energy harvesting networks and statistical signal processing. He is serving as an Editor for *IEEE Transactions on Communications*, *IEEE Transactions on Vehicular Technology*, and *Journal of Wireless Communications and Mobile Computing*, and was an Editor for *IEEE Wireless Communication Letters*, *IEEE Communication Letters* from 2013 to 2016. He received the best paper award in IET Comm. Conf. on Wireless, Mobile and Computing, 2009, IEEE Communication Letter Exemplary Reviewer 2012, and the EU Marie Curie Fellowship 2012-2014.



**Yonghui Li** (M'04-SM'09) received his PhD degree in November 2002 from Beijing University of Aeronautics and Astronautics. From 1999 - 2003, he was affiliated with Linkair Communication Inc, where he held a position of project manager with responsibility for the design of physical layer solutions for the LAS-CDMA system. Since 2003, he has been with the Center of Excellence in Telecommunications, the University of Sydney, Australia. He is now a Professor in School of Electrical and Information Engineering, University of Sydney. He is the recipient of the Australian Queen Elizabeth II Fellowship in 2008 and the Australian Future Fellowship in 2012.

His current research interests are in the area of wireless communications, with a particular focus on MIMO, millimeter wave communications, machine to machine communications, coding techniques and cooperative communications. He holds a number of patents granted and pending in these fields. He is now an editor for IEEE transactions on communications and IEEE transactions on vehicular technology. He received the best paper awards from IEEE International Conference on Communications (ICC) 2014 and IEEE Wireless Days Conferences (WD) 2014.



**Lingyang Song** (S'03-M'06-SM'12) received his PhD from the University of York, UK, in 2007, where he received the K. M. Stott Prize for excellent research. He worked as a research fellow at the University of Oslo, Norway until rejoining Philips Research UK in March 2008. In May 2009, he joined the School of Electronics Engineering and Computer Science, Peking University, China, as a full professor. His main research interests include MIMO, cognitive and cooperative communications, security, and big data. Dr. Song wrote 2 text books, "Wireless

Device-to-Device Communications and Networks" published by Cambridge University Press, UK. He is the recipient of IEEE Leonard G. Abraham Prize in 2016 and IEEE Asia Pacific (AP) Young Researcher Award in 2012. He is currently on the Editorial Board of IEEE Transactions on Wireless Communications. He is an IEEE distinguished lecturer since 2015.