

Energy Efficient Dynamic Resource Optimization in NOMA Systems

Haijun Zhang¹, Senior Member, IEEE, Baobao Wang, Chunxiao Jiang², Senior Member, IEEE,
Keping Long, Senior Member, IEEE, Arumugam Nallanathan³, Fellow, IEEE,
Victor C. M. Leung⁴, Fellow, IEEE, and H. Vincent Poor⁵, Fellow, IEEE

Abstract—Non-orthogonal multiple access (NOMA) with successive interference cancellation (SIC) is a promising technique for next generation wireless communications. Using NOMA, more than one user can access the same frequency-time resource simultaneously and multi-user signals can be separated successfully using SIC. In this paper, resource allocation algorithms for subchannel assignment and power allocation for a downlink NOMA network are investigated. Different from the existing works, here, energy efficient dynamic power allocation in NOMA networks is investigated. This problem is explored using the Lyapunov optimization method by considering the constraints on minimum user quality of service and the maximum transmit power limit. Based on the framework of Lyapunov optimization, the problem of energy efficient optimization can be broken down into three subproblems, two of which are linear and the rest can be solved by introducing a Lagrangian function. The mathematical analysis and simulation results confirm that the proposed scheme can achieve a significant utility performance gain and the energy efficiency and delay tradeoff is derived as $[O(1/V), O(V)]$ with V as a control parameter under maintaining the queue stability.

Index Terms—NOMA, Lyapunov optimization, power allocation, subchannel assignment.

I. INTRODUCTION

IN THE past decade, orthogonal frequency division multiple access (OFDMA) has been widely studied and has been adopted in 4th generation (4G) mobile communication systems [1], [2]. However, orthogonal channel access in OFDMA is becoming a limiting factor of spectral efficiency since each subchannel can only be used by at most one user in each time slot. With the explosive growth of smart mobile devices and the increasing demands for higher spectral efficiency, non-orthogonal multiple access (NOMA) has been proposed to mitigate the consequent heavy loading at the base station (BS) [3]–[5]. NOMA is a promising technique to realize the massive connectivity in 5th generation (5G) mobile networks because NOMA can achieve significant improvement in spectral efficiency with a lower receiver complexity by allowing multiple users to share the same subchannel in the power domain [6]–[8].

The use of NOMA will result in inter-user interference since multiple users will share the same resources [9]. Successive interference cancellation (SIC) can be applied at the end-user receivers to mitigate this interference [10]. NOMA can achieve a capacity region that significantly outperforms orthogonal multiple access schemes by power domain multiplexing at the transmitter and SIC at the receivers [11]. The outage performance of NOMA was evaluated in [12], while in [13], the authors investigated the system sum-rate of multiuser NOMA single-carrier systems in addition to proposing a suboptimal power allocation scheme and presenting a precoder design. Moreover, the authors addressed fairness considerations and posed a max-min fairness problem for NOMA. In [15], an optimal power allocation strategy for the energy-efficiency maximization of NOMA in single-carrier systems was investigated. Fang *et al.* [16] investigated subchannel assignment and power allocation using the difference of convex programming method. A suboptimal algorithm was proposed to solve an uplink scheduling problem with fixed transmission power for uplink NOMA in [17] and a greedy-based algorithm was proposed to improve the throughput in uplink NOMA in [18]. In [19], a new spectrum and energy efficient millimeter-wave transmission scheme which integrates the concept of NOMA with beamspace multiple-input multiple-output (MIMO) was proposed to break the fundamental limit. In [20], a hierarchical power control solution to improve the spectral and energy efficiency in NOMA-enabled vehicular

Manuscript received May 20, 2017; revised October 4, 2017, January 17, 2018, and May 9, 2018; accepted May 22, 2018. Date of publication June 20, 2018; date of current version September 10, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61471025, Grant 61771044, and Grant 61370191, in part by the Young Elite Scientist Sponsorship Program by CAST under Grant 2016QNR001, in part by the Research Foundation of Ministry of Education of China and China Mobile under Grant MCM20170108, in part by the Beijing Natural Science Foundation under Grant L172025, in part by the Fundamental Research Funds for the Central Universities under Grant FRF-GF-17-A6, and in part by the U.S. National Science Foundation under Grants CNS-1702808 and ECCS-1647198. This paper was presented in part at the IEEE Global Communications Conference, Singapore, December 2017. The associate editor coordinating the review of this paper and approving it for publication was L. Duan. (Corresponding authors: Haijun Zhang; Keping Long.)

H. Zhang and K. Long are with the Beijing Engineering and Technology Research Center for Convergence Networks and Ubiquitous Services, University of Science and Technology Beijing, Beijing 100083, China (e-mail: haijunzhang@ieee.org; longkeping@ustb.edu.cn).

B. Wang is with the College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China (e-mail: eeabaowang@gmail.com).

C. Jiang is with the Tsinghua Space Center, Tsinghua University, Beijing 100084, China (e-mail: chx.jiang@gmail.com).

A. Nallanathan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mail: nallanathan@ieee.org).

V. C. M. Leung is with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: vleung@ece.ubc.ca).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TWC.2018.2844359

small cell networks was proposed by performing the joint optimization of cell association and power control. In [21], a novel resource allocation design was investigated for NOMA enhanced heterogeneous networks. However, the problem of energy efficient subchannel assignment and power allocation in multiple small cells with NOMA under the constraints of minimum QoS requirements and cross interference has not been well studied.

In contrast to the typically full buffer assumptions and snapshot-based models, delay is a key metric to measure the QoS. The congestion performance for delay-tolerant services and the stochastic and time-varying features of traffic arrivals should be considered in realistic wireless networks. By utilizing Lyapunov optimization, which is a useful method for handling queue-aware radio resource allocation problems, [22] and [23] considered the impacts of stochastic traffic arrivals and time-varying channel conditions on the system performance to stabilize the queues of networks when optimizing performance metrics. Li *et al.* [22] applied the framework of Lyapunov optimization to balance the average throughput and average delay in energy efficient OFDMA heterogeneous cloud networks. Yu *et al.* [23] adopted the Lyapunov optimization framework in a two-tier OFDMA heterogeneous network under the hybrid access mode to solve the dynamic optimization problem of resource allocation. In [24], an energy efficient resource allocation algorithm was proposed to provide fairness among different small cell base stations (SCBSs) based on the Nash bargaining solution. However, most of the existing works have considered resource allocation using Lyapunov optimization only in OFDMA systems. And to the best of the authors' knowledge, energy efficient resource allocation for time-varying NOMA networks has not been well studied in the previous works.

In this paper, we investigate the subchannel and power allocation respectively in a multiple downlink NOMA network by considering energy efficiency, QoS requirements, power limits, and queue stability. Reference [25] is a conference version of this paper. Different from the conference version, this paper extends [25] to multiple base station scenarios with the detailed proof for the theorem and resource optimization algorithm. We also consider the complexity analysis for the proposed algorithms and more simulation results are provided to verify the proposed methods in this paper compared to the conference version. The main contributions of this paper are summarized as follows:

- Development of a novel energy efficient NOMA network optimization framework: This framework jointly considers energy efficiency maximization, QoS requirements, queue stability and power limits in the optimization of NOMA. Moreover, the queue stability, utility and average queue length performance are studied through both analytical and simulation results.
- Design of a subchannel assignment algorithm based on matching theory: We model the subchannel-user matching problem as a two-sided matching process and propose a subchannel assignment algorithm based on matching.
- Design of a power allocation algorithm with multiple constraints: We formulate a power allocation problem for

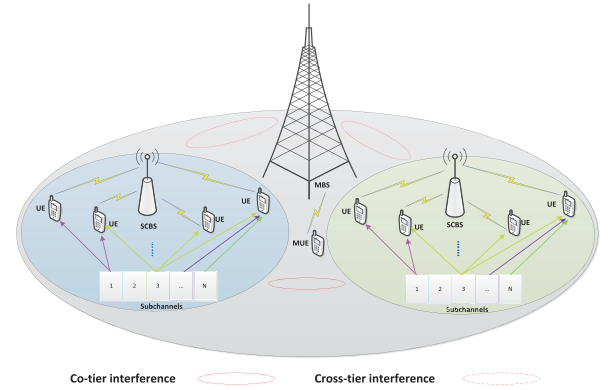


Fig. 1. The architecture of the NOMA network.

NOMA as a mixed integer programming problem. A minimum QoS requirement is employed to provide reliable transmission for users. The energy efficiency optimization problem is decomposed into three subproblems with time-averaged variables and instantaneous variables. We solve the three subproblems and propose a power allocation algorithm using Lyapunov optimization, where the convergence of the proposed algorithm is also demonstrated via simulations.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. Basic Notation

As shown in Fig. 1, we consider a time-varying downlink multiple cell NOMA network in which N SCBSs exist and each SCBS transmits the signals to each set of mobile users denoted by $\mathcal{U} = \{1, 2, \dots, U\}$. The available bandwidth is divided by the BS into a set of subchannels which is denoted by $\mathcal{N} = \{1, 2, \dots, N\}$. The NOMA system is assumed to operate in a slotted time mode with unit time slots $t \in \{0, 1, 2, \dots\}$, where the time slot t refers to the time interval $[t, t+1)$. We assume that the BSs have full knowledge of the channel state information and denote by $g_{k,u,n}(t)$ the channel gain between the k th SCBS and user u on subchannel n at time slot t . We set $a_{k,u,n}(t) = 1$ when the subchannel n of SCBS k is allocated to user u at time slot t ; otherwise, $a_{k,u,n}(t) = 0$. $[x]^+$ denotes the larger of x and zero, x^T denotes the transpose of x , and $E\{\cdot\}$ denotes the expectation.

B. System Model

In NOMA, one user can receive signals from the BS through multiple subchannels and one subchannel can be allocated to multiple users at the same time slot t . Since the user j on subchannel n causes interference to the other users on the same subchannel, each user j adopts SIC after receiving the superposed signals to demodulate the target message. As shown in [26], without constraints on the specific power split, one user's data can be successfully decoded by another user whose channel gain is better via superposition coding with SIC. Since user u with higher channel gain can only decode the signals of user i with worse channel gain, the interference signals caused by user j whose channel gain is better than user u cannot be decoded and will be treated as noise. Thus,

after SIC, the interference for user u caused by other users of same SCBS k on the same subchannel n is given by

$$I_{k,u,n}(t) = \sum_{i \in \{S_{k,n} | g_{k,i,n} > g_{k,u,n}\}} a_{k,i,n}(t) p_{k,i,n}(t) g_{k,u,n}(t) \quad (1)$$

where $S_{k,n}$ is the set of users of SCBS k on subchannel n . Modeling this residual interference as additional AWGN, we can use the Shannon's capacity formula to write the capacity of user $u \in \mathcal{U} = \{1, 2, \dots, U\}$ of SCBS k on the n th subchannel at time slot t as

$$R_{k,u,n}(t) = \frac{B}{N} a_{k,u,n}(t) \log_2 \left(1 + \frac{p_{k,u,n}(t) g_{k,u,n}(t)}{\sigma_{k,n}^2 + I_{k,u,n}(t) + \tilde{I}_{k,u,n}(t)} \right) \quad \forall k \in K, u \in U, n \in N, \quad (2)$$

where B and σ_n^2 are the bandwidth of the system and the noise variance respectively. $\tilde{I}_{k,u,n}(t)$ is the total of the co-tier interference caused by other SCBSs to SCBS k and cross-tier interference caused by the macro BS to SCBS k which is given by

$$\tilde{I}_{k,u,n}(t) = \sum_{l \neq k} \sum_{u=1}^U a_{l,u,n} p_{l,u,n} g_{l,k,u,n} + \sum_{u=1}^U p_{u,n}^M g_{k,u,n}^M \quad \forall k \in K, u \in U, n \in N, \quad (3)$$

where $g_{l,k,u,n}$ is the channel gain on subchannel n of user u in SCBS l to SCBS k . $p_{u,n}^M$ and $g_{k,u,n}^M$ are the power allocation and channel gain on subchannel n of user u in the macro BS to SCBS k respectively. The capacity of user u of SCBS k at time slot t can be written as

$$R_{k,u}(t) = \sum_{n=1}^N R_{k,u,n}(t), \quad \forall k \in K, u \in U. \quad (4)$$

In order to specify the QoS of users, we let $\hat{R}_{k,u}$ be the QoS requirement in terms of the minimum capacity of user u which is thus given as

$$C1: R_{k,u}(t) \geq \hat{R}_{k,u}, \quad \forall k \in K, u \in U. \quad (5)$$

When the queue stability of the SCBS is guaranteed, we can focus only on the performance optimization with queue stability. This is because the SCBS can be assumed to be fixed over a longer duration than the scheduling slot of subchannels. In this NOMA network, the separate buffering queue $Q_{k,u}(t)$ is maintained for each user U of SCBS k . The random traffic arrivals for $Q_{k,u}(t)$ are denoted by $A_{k,u}(t)$ whose peak arrival is $A_{k,u}^{\max}$ at time slot t . $A_{k,u}(t)$ is independent and identically distributed (i.i.d.) over slots. The exogenous arrival rates may be outside of the network capacity region in practice for the reasons that the statistic of $A_{k,u}(t)$ is usually unknown to SCBSs and the achievable capacity region is usually difficult to estimate. Then, a transport flow control mechanism is needed to keep the traffic queues stabilized. Denote by $r_{k,u}(t)$ the admitted data rate out of the potentially substantial traffic arrivals for user u of SCBS k which obviously satisfies

$0 \leq r_{k,u}(t) \leq A_{k,u}(t)$. Therefore, we denote the traffic buffering queues for user u as

$$Q_{k,u}(t+1) = [Q_{k,u}(t) - R_{k,u}(t)]^+ + r_{k,u}(t), \quad \forall k \in K, u \in U. \quad (6)$$

The time averaged throughput in terms of the data arrivals for users in SCBS k is defined as $\bar{r}_{k,u} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T r_{k,u}(t)$.

Denote by $p_{k,u}(t)$ and $p_{tot}(t)$ the instantaneous power of user u of SCBS k at time slot t and the total power consumption of all the SCBSs at time slot t respectively, which can be written as

$$p_{k,u}(t) = \sum_{n=1}^N p_{k,u,n}(t) + p_k^C \quad (7)$$

and

$$p_{tot}(t) = \sum_{k=1}^K \sum_{u=1}^U p_{k,u}(t) \quad (8)$$

where p_k^C accounts for the circuit power of SCBS k . The average and instantaneous power constraints of user u of SCBS k are denoted by $P_{k,u}$ and $\hat{P}_{k,u}$, which can be written as

$$\bar{p}_{k,u} = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=0}^t p_{k,u}(\tau) \leq P_{k,u}, \quad \forall k \in K, u \in U \quad (9)$$

$$p_{k,u}(t) \leq \hat{P}_{k,u}, \quad \forall k \in K, u \in U. \quad (10)$$

We define the function $g_R(\cdot)$ as the revenue obtained by the throughput which is a non-decreasing concave utility function. Let η_{EE} denote the energy efficiency (EE) which is defined as the ratio of the profit brought by the long-term utility of average throughput to the corresponding long-term total power consumption. It can be written as

$$\eta_{EE} = \frac{\sum_{k=1}^K \sum_{u=1}^U g_R(\bar{r}_{k,u})}{\bar{p}_{tot}}, \quad (11)$$

where $\bar{p}_{tot} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T p_{tot}(t)$.

C. Optimization Problem Formulation

In this subsection, when considering all constraints, the utility function is expressed as

$$\begin{aligned} \max_{\mathcal{P}} \eta_{EE} &= \frac{\sum_{k=1}^K \sum_{u=1}^U g_R(\bar{r}_{k,u})}{\bar{p}_{tot}} \\ s.t. \quad C1 \quad C2: \quad &\bar{p}_{k,u} = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=0}^t p_{k,u}(\tau) \leq P_{k,u}, \\ &\forall k \in K, u \in U \\ C3: \quad &p_{k,u}(t) \leq \hat{P}_{k,u}, \quad \forall k \in K, u \in U \\ C4: \quad &\widetilde{R_{k,u}} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T R_{k,u}(t) \geq \bar{r}_{k,u}, \\ &\forall k \in K, u \in U. \end{aligned} \quad (12)$$

where $\mathcal{P}$ denotes the power allocation policy and the constraint $C1$ ensures the QoS of users; $C2$ is a bound on the limit of the maximum average transmit power of user u of SCBS k ; $C3$ is the maximum instantaneous transmit power of user u of SCBS k ; and $C4$ ensures the stability of user u in SCBS k . We define the optimal EE η_{EE}^{opt} as

$$\eta_{EE}^{opt} = \frac{\sum_{k=1}^K \sum_{u=1}^U g_R(\bar{r}_{k,u}(\mathcal{P}^*))}{\bar{p}_{tot}(\mathcal{P}^*)} = \max_{\mathcal{P}} \frac{\sum_{k=1}^K \sum_{u=1}^U g_R(\bar{r}_{k,u}(\mathcal{P}))}{\bar{p}_{tot}(\mathcal{P})} \quad (13)$$

where $\mathcal{P}^*$ denotes the optimal power allocation policy that yields η_{EE}^{opt} . We can classify the utility function in (12) as a nonlinear fractional program. We introduce Theorem 1 as follows.

Theorem 1: The optimal EE η_{EE}^{opt} can be reached if and only if

$$\begin{aligned} & \max_{\mathcal{P}} \sum_{k=1}^K \sum_{u=1}^U g_R(\bar{r}_{k,u}(\mathcal{P})) - \eta_{EE}^{opt} \bar{p}_{tot}(\mathcal{P}) \\ &= \sum_{k=1}^K \sum_{u=1}^U g_R(\bar{r}_{k,u}(\mathcal{P}^*)) - \eta_{EE}^{opt} \bar{p}_{tot}(\mathcal{P}^*) = 0 \\ & \text{for } \sum_{k=1}^K \sum_{u=1}^U g_R(\bar{r}_{k,u}(\mathcal{P})) \geq 0, \bar{p}_{tot}(\mathcal{P}) \geq 0. \end{aligned} \quad (14)$$

Proof: Please refer to Appendix A.

According to Theorem 1, an equivalent objective function in subtractive form exists for the (non-convex) optimization problem (12) which can be classified as a nonlinear fractional program. Then, the formulation (12) can be rewritten in the more tractable form

$$\begin{aligned} & \max_{\mathcal{P}} \sum_{k=1}^K \sum_{u=1}^U g_R(\bar{r}_{k,u}(\mathcal{P})) - \eta_{EE} \bar{p}_{tot}(\mathcal{P}) \\ & \text{s.t. } C1, C2, C3, C4. \end{aligned} \quad (15)$$

III. ENERGY EFFICIENT OPTIMIZATION USING LYAPUNOV OPTIMIZATION

In this section, subchannel assignment is investigated in the NOMA network and the optimization problem in (15) is solved based on Lyapunov optimization.

A. Subchannel Matching

We assume that all the users of BS can transmit on the subchannel n arbitrarily at time slot t in a NOMA system. Considering the complexity of decoding and the fairness of users, each subchannel can only be allocated to at most D_n users and each user can only occupy at most D_u subchannels at the same time. We assume that $N * D_n \geq U * D_u$. The dynamic matching between the users and the subchannels of SCBS k can be considered as a many-many matching process between the set of U users and the set of N subchannels. User u is matched with subchannel n at time slot t if $a_{k,u,n}(t) = 1$. Based on the channel state information, we assume user u prefers channel n_1 over n_2 if and only if $g_{k,u,n_1} > g_{k,u,n_2}$.

Algorithm 1 Suboptimal Matching Algorithm for Subchannel Allocation

- 1: Initialize the matched lists $S_{k,n}$ and $S_{k,u}$ to denote the number of users matched with subchannel n ($\forall n \in \{1, 2, \dots, N\}$) and the number of subchannels matched with user u ($\forall u \in \{1, 2, \dots, U\}$) in SCBS k , respectively;
- 2: Initialize the preference lists $Pref_U(k, u)$ for all the users of SCBS k according to channel state information;
- 3: Initialize the set of not fully matched users $S_{U_F}(k, u)$ to denote users of SCBS k who have not been matched with D_u subchannels;
- 4: **while** $S_{U_F}(k, u) \neq \emptyset$ **do**
- 5: **for** $u = 1$ to U **do**
- 6: **if** $S_{k,u} < D_u$ **then**
- 7: User u of SCBS k sends a matching request to its most preferred subchannel $\hat{n}$ according to $Pref_U(k, u)$;
- 8: **if** $S_{k,\hat{n}} < D_{k,\hat{n}}$ **then**
- 9: Set $a_{k,u,\hat{n}} = 1$, $S_{k,u} = S_{k,u} + 1$ and $S_{k,\hat{n}} = S_{k,\hat{n}} + 1$;
- 10: **else if** $S_{k,\hat{n}} = D_{k,\hat{n}}$ **then**
- 11: Find the minimum channel gain of users $g_{k,\hat{u},\hat{n}}$ on channel $\hat{n}$ in SCBS k and compare it with $g_{k,u,\hat{n}}$;
- 12: **if** $g_{k,\hat{u},\hat{n}} < g_{k,u,\hat{n}}$ **then**
- 13: Set $a_{k,u,\hat{n}} = 1$, $a_{k,\hat{u},\hat{n}} = 0$, $S_{k,u} = S_{k,u} + 1$, and $S_{k,\hat{u}} = S_{k,\hat{u}} - 1$;
- 14: **else**
- 15: Remove subchannel $\hat{n}$ from $Pref_U(k, u)$ and find the next $\hat{n}$ of user u in SCBS k according to $Pref_U(k, u)$.
- 16: **end if**
- 17: **end if**
- 18: **end for**
- 19: **end while**

Then, the preference lists of the users of SCBS k can be written as

$$Pref_U(K, U) = [Pref_U(1), \dots, Pref_U(u), \dots, Pref_U(U)]^T \quad (16)$$

where $Pref_U(K, U)(u)$ is the preference list of user u of SCBS k which is in the descending order of channel gains of subchannels. To reduce the complexity, we propose a suboptimal matching algorithm for subchannel allocation as Algorithm 1. For each user u of SCBS k , the matching request is sent to its most preferred subchannel according to its preference list. Then the preferred subchannel decides whether to accept the user or not according to the users's channel gain. For each subchannel n , there are at most D_n accepted users who have higher channel gains than the rejected users. A rejected user on subchannel n will remove the subchannel n from its preference list and its most preferred subchannel is changed. The matching request is send again until all the users of each SCBS have matched D_u subchannels and the matching is achieved.

B. The Queues of Lyapunov Optimization

Because $g_R(\bar{r}_{k,u})$ is related to the time averaged throughput, we define auxiliary variables $\gamma_{k,u}$ for the traffic arrival of user u in SCBS k which satisfies $\bar{\gamma}_{k,u} \leq \bar{r}_{k,u}$ and $0 \leq \gamma_{k,u} \leq A_{k,u}^{\max}$. Therefore, the optimization problem in (15) can be rewritten as

$$\begin{aligned} \max_{\mathcal{P}} \quad & \sum_{k \in K} \sum_{u \in U} \overline{g_R(\gamma_{k,u}) - \eta_{EE} p_{tot}} \\ \text{s.t.} \quad & C1, C2, C3, C4 \\ & C5: \bar{\gamma}_{k,u} \leq \bar{r}_{k,u}, \quad 0 \leq \gamma_{k,u} \leq A_{k,u}^{\max} \end{aligned} \quad (17)$$

where $\bar{\gamma}_{k,u} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \gamma_{k,u}(t)$ and $\overline{g_R(\gamma_{k,u}) - \eta_{EE} p_{tot}} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T (g_R(\gamma_{k,u}(t)) - \eta_{EE} p_{tot})$. The equivalent of problems (15) and (17) is proved in Appendix B. To satisfy the averaged throughput constraint in C5, we denote by $H_{k,u}(t)$ the virtual queue for user u in SCBS k at time slot t . We get

$$\begin{aligned} H_{k,u}(t+1) = [H_{k,u}(t) - r_{k,u}(t)]^+ \\ + \gamma_{k,u}(t), \quad \forall k \in \mathcal{K}, u \in \mathcal{U}. \end{aligned} \quad (18)$$

Similarly, in order to satisfy the constraint C2, virtual power queues for the user u in SCBS k are defined. We denote $Z_{k,u}(t)$ as the queue arrival with the transmit power $p_{k,u}(t)$, and get

$$Z_{k,u}(t+1) = [Z_{k,u}(t) - P_{k,u}]^+ + p_{k,u}(t), \quad \forall k \in \mathcal{K}, u \in \mathcal{U}. \quad (19)$$

C. The Formulation of Lyapunov Optimization

Denote by $\Phi(t) = [Q(t), H(t), Z(t)]$ the matrix of all the queues. We define the Lyapunov function as a scalar metric of queue congestion:

$$L(\Phi(t)) = \frac{1}{2} \left\{ \sum_{k=1}^K \sum_{u=1}^U (Q_{k,u}(t)^2 + H_{k,u}(t)^2 + Z_{k,u}(t)^2) \right\}. \quad (20)$$

We introduce a Lyapunov drift in this subsection for pushing the Lyapunov function to a lower congestion state and keep both the actual and virtual queues stable:

$$\Delta(\Phi(t)) = E \{ L(\Phi(t+1)) - L(\Phi(t)) \}. \quad (21)$$

According to Lyapunov optimization, by subtracting $VE \left\{ \sum_{u \in \mathcal{U}} g_R(\gamma_{k,u}) - \eta_{EE} p_{tot} \right\}$ from both sides of (21),

we get

$$\begin{aligned} \Delta(\Phi(t)) - VE \left\{ \sum_{k=1}^K \sum_{u=1}^U g_R(\gamma_{k,u}) - \eta_{EE} p_{tot} \right\} \\ \leq C + \sum_{k=1}^K \sum_{u=1}^U E \{ H_{k,u}(t) \gamma_{k,u} - V g_R(\gamma_{k,u}) \} \\ + VE \left(\eta_{EE} \left\{ \sum_{k=1}^K \left\{ \sum_{u=1}^U p_{k,u}(t) + p_k^C \right\} \right\} \right) \\ - \sum_{k=1}^K \sum_{u=1}^U \{ H_{k,u}(t) - Q_{k,u}(t) \} E \{ r_{k,u}(t) \} \\ - \sum_{k=1}^K \sum_{u=1}^U Q_{k,u}(t) E \{ R_{k,u}(t) \} \\ - \sum_{k=1}^K \sum_{u=1}^U Z_{k,u}(t) E \{ P_{k,u}(t) - p_{k,u}(t) \} \end{aligned} \quad (22)$$

where V is an arbitrarily positive control parameter which represents the emphasis on utility maximization compared to queue stability and C is a finite constant that satisfies

$$\begin{aligned} C \geq \frac{1}{2} E \left\{ \sum_{k=1}^K \sum_{u=1}^U (r_{k,u}(t)^2 + R_{k,u}(t)^2) \right\} \\ + \frac{1}{2} E \left\{ \sum_{k=1}^K \sum_{u=1}^U [(r_{k,u}(t) - \gamma_{k,u}(t))^2 \right. \\ \left. + (P_{k,u}(t) - p_{k,u}(t))^2] \right\}. \end{aligned} \quad (23)$$

1) *The Solution of Virtual Variables:* The optimal choice of γ_u to minimize (22) can be made by solving

$$\begin{aligned} \max V g_R(\gamma_{k,u}) - H_{k,u}(t) \gamma_{k,u} \\ \text{s.t. } 0 \leq \gamma_{k,u} \leq A_{k,u}^{\max} \end{aligned} \quad (24)$$

and the solution is

$$\gamma_{k,u}(t) = \min \left\{ \frac{V}{H_{k,u}(t)}, A_{k,u}^{\max} \right\}. \quad (25)$$

2) *The Solution of Actual Traffic Arrival:* In order to minimize (22), the optimal actual traffic arrival can be achieved by maximizing the expression as follows:

$$\begin{aligned} \max \{ H_{k,u}(t) - Q_{k,u}(t) \} E \{ r_{k,u}(t) \} \\ \text{s.t. } 0 \leq r_{k,u}(t) \leq A_{k,u}(t). \end{aligned} \quad (26)$$

We get the optimal solution as

$$r_{k,u}(t) = \begin{cases} A_{k,u}(t), & \text{if } H_{k,u}(t) - Q_{k,u}(t) \succ 0 \\ 0, & \text{else.} \end{cases} \quad (27)$$

3) *Power Allocations:* In order to minimize (22), we first obtain the optimal solution of the virtual variables and the actual traffic arrival, then we can minimize the remaining part

$$\begin{aligned}
\min V E \left(\eta_{EE} \left\{ \sum_{k=1}^K \sum_{u=1}^U p_{k,u}(t) \right\} \right) & - \sum_{k=1}^K \left\{ \sum_{u=1}^U Q_{k,u}(t) E \{ R_{k,u}(t) \} + \sum_{u=1}^U Z_{k,u}(t) p_{k,u}(t) \right\} \\
& = V E \left(\eta_{EE} \left\{ \sum_{k=1}^K \sum_{u=1}^U \sum_{n=1}^N \omega_{k,u,n}(t) \right\} \right) + \sum_{k=1}^K \sum_{u=1}^U Z_{k,u}(t) \left\{ \sum_{n=1}^N \omega_{k,u,n}(t) \right\} \\
& - \sum_{k=1}^K \sum_{u=1}^U Q_{k,u}(t) E \left\{ \frac{B}{N} \sum_{n=1}^N a_{k,u,n}(t) \log_2 \left(1 + \frac{\omega_{k,u,n}(t) g_{k,u,n}(t)}{a_{k,u,n}(t) (\sigma_{k,n}^2 + I_{k,u,n}(t) + \tilde{I}_{k,u,n}(t))} \right) \right\} \quad (30)
\end{aligned}$$

$$\begin{aligned}
F(\lambda, \beta) = \min L(\lambda, \beta) & = V E \left(\eta_{EE} \left\{ \sum_{k=1}^K \sum_{u=1}^U \sum_{n=1}^N \omega_{k,u,n}(t) \right\} \right) + \sum_{k=1}^K \sum_{u=1}^U Z_{k,u}(t) \left\{ \sum_{n=1}^N \omega_{k,u,n}(t) \right\} \\
& - \sum_{k=1}^K \sum_{u=1}^U Q_{k,u}(t) E \left\{ \sum_{n=1}^N \frac{B}{N} a_{k,u,n}(t) \log_2 \left(1 + \frac{\omega_{k,u,n}(t) g_{k,u,n}(t)}{a_{k,u,n}(t) (\sigma_{k,n}^2 + I_{k,u,n}(t) + \tilde{I}_{k,u,n}(t))} \right) \right\} \\
& + \sum_{k=1}^K \sum_{u=1}^U \lambda_{k,u}(t) \left\{ \sum_{n=1}^N \omega_{k,u,n}(t) - \hat{P}_{k,u} \right\} \\
& + \sum_{k=1}^K \sum_{u=1}^U \beta_{k,u}(t) \left\{ \hat{R}_{k,u} - \sum_{n=1}^N \frac{B}{N} a_{k,u,n}(t) \log_2 \left(1 + \frac{\omega_{k,u,n}(t) g_{k,u,n}(t)}{a_{k,u,n}(t) (\sigma_{k,n}^2 + I_{k,u,n}(t) + \tilde{I}_{k,u,n}(t))} \right) \right\} \quad (31)
\end{aligned}$$

of (22) and denote it as

$$\begin{aligned}
\min & \left\{ V E \left(\eta_{EE} \left\{ \sum_{k=1}^K \left\{ \sum_{u=1}^U p_{k,u}(t) + p_k^C \right\} \right\} \right) \right. \\
& + \sum_{k=1}^K \sum_{u=1}^U Z_{k,u}(t) p_{k,u}(t) \\
& \left. - \sum_{k=1}^K \sum_{u=1}^U Q_{k,u}(t) E \{ R_{k,u}(t) \} \right\} \\
& = V E \left(\eta_{EE} \left\{ \sum_{k=1}^K \left\{ \sum_{u=1}^U \sum_{n=1}^N a_{k,u,n}(t) p_{k,u,n}(t) \right\} + p_k^C \right\} \right) \\
& + \sum_{k=1}^K \sum_{u=1}^U Z_{k,u}(t) \left\{ \sum_{n=1}^N a_{k,u,n}(t) p_{k,u,n}(t) \right\} \\
& - \sum_{k=1}^K \sum_{u=1}^U Q_{k,u}(t) E \left\{ \sum_{n=1}^N R_{k,u,n}(t) \right\}. \quad (28)
\end{aligned}$$

Let $\omega_{k,u,n}(t) = a_{k,u,n}(t) p_{k,u,n}(t)$, $\forall k \in \mathcal{K}, u \in \mathcal{U}, n \in \mathcal{N}$; then we can get

$$\begin{aligned}
R_{k,u,n}(t) & = \frac{B}{N} a_{k,u,n}(t) \log_2 \\
& \times \left(1 + \frac{\omega_{k,u,n}(t) g_{k,u,n}(t)}{a_{k,u,n}(t) (\sigma_{k,n}^2 + I_{k,u,n}(t) + \tilde{I}_{k,u,n}(t))} \right). \quad (29)
\end{aligned}$$

We can rewrite (28) as (30), shown at the top of this page. Since (30) is convex, to satisfy the series of constraints, the Lagrangian of the problem (30) can be expressed by (31),

shown at the top of this page, where λ and β are the Lagrange multiplier vectors for the constraints in (17). Taking the first order derivation of $F(\lambda, \beta)$ with respect to $\omega_{k,u,n}(t)$, we can get the optimal power allocation as follows:

$$\begin{aligned}
p_{k,u,n}(t) & = \frac{\omega_{k,u,n}(t)}{a_{k,u,n}(t)} \\
& = \frac{\frac{B}{N} (Q_{k,u}(t) + \beta_{k,u}(t))}{\ln 2 (V \eta_{EE} + Z_{k,u}(t) + \lambda_{k,u}(t))} \\
& - \frac{\sigma_{k,n}^2 + I_{k,u,n}(t) + \tilde{I}_{k,u,n}(t)}{g_{k,u,n}(t)}. \quad (32)
\end{aligned}$$

Based on the subgradient method [28], the master dual problem in (32) can be solved by

$$\begin{aligned}
\lambda_{k,u}^{l+1} & = \left[\lambda_{k,u}^l - \varepsilon_1^l \left(\hat{P}_{k,u} - \sum_{n=1}^N \omega_{k,u,n}(t) \right) \right]^+, \quad \forall k \in \mathcal{K}, u \in \mathcal{U}; \\
\beta_{k,u}^{l+1} & = \left[\beta_{k,u}^l - \varepsilon_2^l \left(\sum_{n=1}^N R_{k,u,n} - R_{k,u} \right) \right]^+, \quad \forall k \in \mathcal{K}, u \in \mathcal{U}. \quad (33)
\end{aligned}$$

After the time interval T , we get the average queue length $\bar{Q}$, total average capacity R^{ave} , average power consumption P^{ave} and the average energy efficient EE η_{EE}^{ave} as

$$\bar{Q} = \frac{1}{T} \sum_{t=1}^T \sum_{k=1}^K \sum_{u=1}^U Q_{k,u}(t) \quad (34)$$

$$R^{ave} = \frac{1}{T} \sum_{t=1}^T U_{tot}(t) \quad (35)$$

$$\bar{P} = \frac{1}{T} \sum_{t=1}^T \sum_{k=1}^K \sum_{u=1}^U p_{k,u}(t) \quad (36)$$

$$\eta_{EE}^{ave} = \frac{1}{T} \sum_{t=1}^T \eta_{EE}. \quad (37)$$

The above proposed approach based on Lyapunov optimization for solving the EE optimization problem in (19) can be summarized in Algorithm 2.

Algorithm 2 Lyapunov Optimization Based Resource Allocation Algorithm

- Initialize the $a_{k,u,n}(t)$ using suboptimal Algorithm 1;
 2: Initialize $p_{k,u,n}(t)$ using equal power allocation;
 Initialize the value of $A_{k,u}^{\max}$ and $Q_{k,u}(t)$;
 4: For each time slot, calculate the auxiliary variables $\gamma_{k,u}(t)$ for each time slot and admitted traffic $r_{k,u}(t)$ by solving (25) and (27) respectively;
repeat
 6: Obtain the optimal allocation of power in the current time slot according to (32);
 Update the Lagrange multipliers λ and β by solving (33);
 8: Solve (12) using the Dinkelbach method in [1] to get the optimal value of η_{EE} in an iteration way;
until Convergence or certain stopping criteria is met
 10: Calculate the traffic queue $Q_{k,u}(t)$ and the virtual queues of $H_{k,u}(t)$ and $Z_{k,u}(t)$ of next time slot by solving (6), (18), (19) respectively.
-

D. Complexity Analysis

The complexity of the proposed algorithms is analyzed in this subsection. The worst-case complexity of proposed suboptimal Algorithm 1 is $O(KUN + 1/2K[U/D_n]^2)$ in which the finding of preference lists is $O(KUN)$. The optimal algorithm of subchannel matching can only be obtained by searching over all possible combinations of users whose computational complexity can be approximated as $O(K \frac{[U]^2}{D_n^2})$. We see from the above analysis that the proposed suboptimal subchannel matching Algorithm 1 has a much lower polynomial complexity than exhaustive search. In Algorithm 2, the calculation of (32) for each user in each SCBS on each subchannel entails KUN operations. Thus the complexity of Algorithm 2 is $O(TLKUN)$ where L is the number of iterations in each time slot. As the complexity of Algorithm 2 increases with the number of users and the convergence of EE is influenced by the control parameter V , Algorithm 2 may be practical for a middle-scale realtime network and the value of V should be well chosen. Furthermore, cellular users are typically distributed in clusters. Therefore, the cellular users in a cell can be divided into several clusters, and the proposed algorithm can be used in each cluster. Subchannels are not shared between different clusters. Since the number of users in each cluster is small, the complexity of the proposed algorithms will be further reduced.

IV. PERFORMANCE ANALYSIS

In this section, we analyze the performance bounds of the proposed Algorithm 2 based on Lyapunov optimization.

A. Stability of Queues

In this work, all the actual queues $Q_{k,u}(t)$ and virtual queues of $H_{k,u}(t)$ and $Z_{k,u}(t)$ are mean rate stable, as we will now show. We assume the expectation of $P_{tot}(t)$ and $U_{tot}(t)$ are bounded as

$$P_{\min} \leq E\{P_{tot}(t)\} \leq P_{\max} \quad (38)$$

$$R_{\min} \leq E\{U_{tot}(t)\} \leq R_{\max} \quad (39)$$

where $P_{\min}$, $P_{\max}$, $R_{\min}$ and $R_{\max}$ are finite constants. From the boundedness assumptions, there is a positive constant C satisfying

$$\Delta(\Phi(t)) \leq C. \quad (40)$$

It can be written as

$$E\{L(\Phi(t+1))\} - E\{L(\Phi(t))\} \leq C. \quad (41)$$

Considering telescoping sums over $t \in \{0, 1, \dots, T-1\}$ in the above inequality, we can get

$$E\{L(\Phi(T))\} - E\{L(\Phi(0))\} \leq TC. \quad (42)$$

Using (20), we can rewrite (42) as

$$E\{Z_{k,u}(T)^2\} \leq 2TC + 2E\{L(\Phi(0))\}. \quad (43)$$

Given $D(|Z_{k,u}(T)|) = E(|Z_{k,u}(T)|^2) - [E(|Z_{k,u}(T)|)]^2 \geq 0$, we get $E(|Z_{k,u}(T)|^2) \geq [E(|Z_{k,u}(T)|)]^2$. Thus we have

$$E(|Z_{k,u}(T)|) \leq \sqrt{2TC + 2E\{L(\Phi(0))\}}. \quad (44)$$

Dividing by T and taking the limit $T \rightarrow \infty$, we have

$$\lim_{T \rightarrow \infty} \frac{E(|Z_{k,u}(T)|)}{T} = 0. \quad (45)$$

Therefore, the queues of $Z_{k,u}(T)$ are mean rate stable. Similarly, we can also prove the queues $Q_{k,u}(T)$ and $H_{k,u}(T)$ are mean rate stable.

B. The Utility and Average Queue Length Performance

The EE performance and the average queue length performance obtained by Algorithm 2 are given in Theorem 2.

Theorem 2: Utilizing the proposed optimization solution, η_{EE} is bounded by

$$\eta_{EE} \geq \eta_{EE}^{opt} - \frac{B}{VP_{\min}}. \quad (46)$$

The performance of the average network queue length is bounded by

$$\bar{Q} \leq \frac{B + V(R_{\max} - \eta_{EE}^{opt}P_{\min})}{\varepsilon} \quad (47)$$

where η_{EE}^{opt} is the theoretical maximum achievable utility of all achievable solutions.

Proof: Please refer to Appendix C.

Since $\eta_{EE} \leq \eta_{EE}^{opt}$, which naturally holds from (13), EE is bounded according to $\eta_{EE}^{opt} - \frac{B}{VP_{\min}} \leq \eta_{EE} \leq \eta_{EE}^{opt}$. Therefore, η_{EE} can arbitrarily approach η_{EE}^{opt} by setting large enough V to make $\frac{B}{VP_{\min}}$ arbitrarily small. The bound on the average queue backlog $\bar{Q}$ increases linearly in V according to (47). Theorem 2 shows that there exists an $[O(1/V), O(V)]$ utility-backlog tradeoff that leads to Little's Theorem.

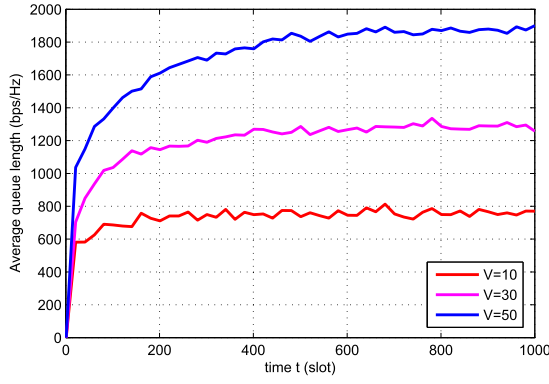


Fig. 2. Average queue length versus simulation time.

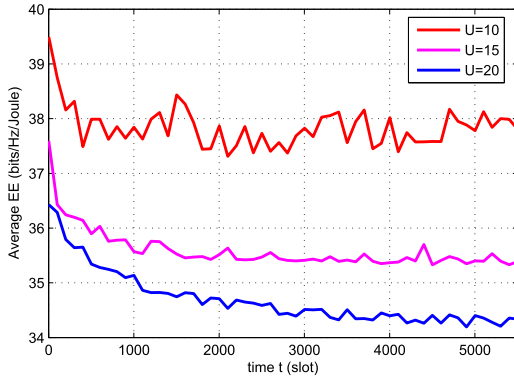


Fig. 3. EE performance versus simulation time.

V. SIMULATION RESULTS AND DISCUSSION

In this section, simulation results are given to evaluate the performance of the proposed algorithms. In the considered NOMA network, there are K SCBSs comprising U users distributed randomly in a circular with radius of 30 m and the SCBSs are distributed randomly in the circular of MBS whose radius is 500 m. The noise power spectral density is -174 dBm/Hz. We consider a normalized bandwidth which is $B = 1$ Hz. For the simulation, the maximum average transmit power and the maximum instantaneous transmit power of user u in SCBS k is set as $2.5/U$ Watts and $2.505/U$ Watts respectively. We assume that the circuit power of user u is 1 Watt. The time interval T is set to 5000 slots which means traffic arrival rates is 15 bps/Hz. We assume that each user can access only one OFDMA subchannel.

In Fig. 2, the average queue length versus time t is evaluated under different values of the control parameter V . The user QoS requirement is set as $\hat{R}_u = 2$ bps/Hz and the maximum number of matched subchannels of each user is $D_u = 2$ while the maximum number of matched users of each subchannel is $D_n = 2$. It is seen that the value of the average queue length increases with time t and gradually fluctuates around a certain fixed value. And for the same value of t , a smaller value of V leads to a smaller value of average queue length which agrees with Theorem 2. This conclusion can also be inferred from Fig. 9.

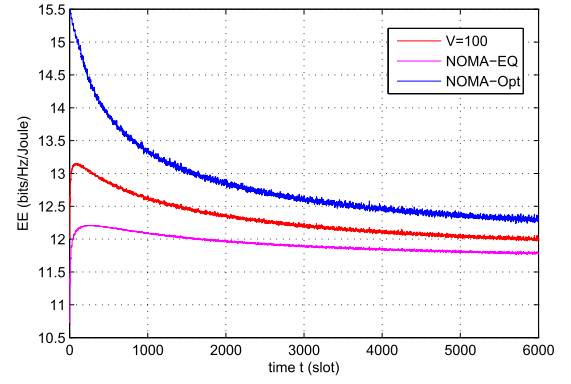


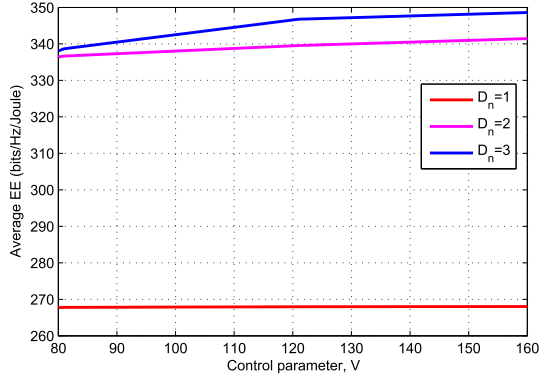
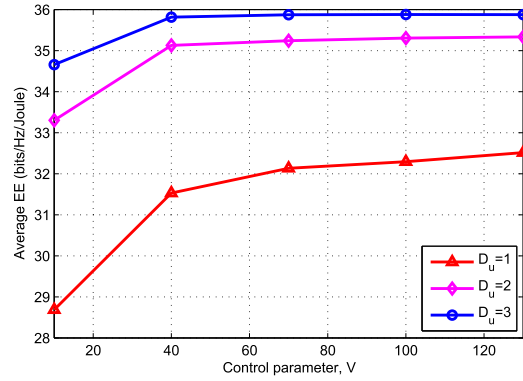
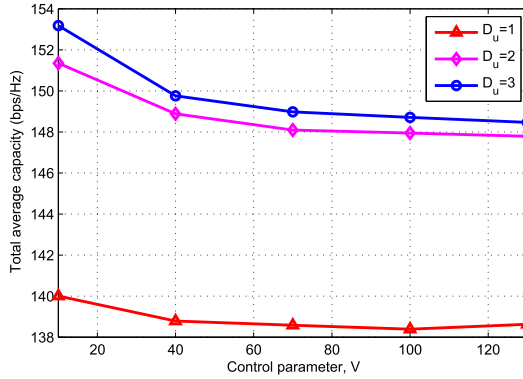
Fig. 4. EE performance versus simulation time for different algorithms.

Fig. 3 shows the EE performance versus time t with the same constraints of Fig. 2 except the QoS constraint is $\hat{R}_u = 1$ bps/Hz instead. It can be observed that the EE performance will gradually fluctuate around a certain fixed value with the increase of t which illustrates the convergence of EE with time t . And larger values of U lead to smaller converged values of EE. This is because the needs of subchannels will increase with an increase in the number of users, which leads to a decrease in the bandwidth of each subchannel.

As shown in Fig. 4, the EE performance is shown versus the time slot t . As t increases, the EE gradually fluctuates around a certain fixed value. From Fig. 4, we observe that the performance of $V = 100$ which is the proposed resource allocation algorithms with the control parameter $V = 100$, including subchannel assignment and power allocation, is much better than the NOMA-EQ algorithm in [16]. This is because the NOMA-EQ algorithm is limited to equal power allocation for each subchannel in NOMA system. For different subchannel allocation schemes, the EE performance using exhaustive search, which is denoted by NOMA-Opt, is better than using our suboptimal algorithm.

In Fig. 5, the average EE performance is shown versus the parameter V with different values of D_n which is the number of users matched with each subchannel and $D_n = 1$ represents OFDMA. The user QoS requirement is set as $\hat{R}_u = 2$ bps/Hz and the maximum number of matched subchannels of each user is $D_u = 1$. It is shown that, with an increase in the parameter V , the value of average EE increases and converges to a certain value both for NOMA at the same value of D_u and for OFDMA. It is seen that the average EE for NOMA is better than the average EE for OFDMA. And for the same value of V , a larger value of D_n leads to a larger value of average EE. This is because our proposed algorithm provides more freedom in the bandwidth allocation of assigned subchannels. For the same set of users under the same value of D_u , the larger of the value of D_n is, the larger is the bandwidth of each subchannel.

Fig. 6 shows the average EE performance versus the parameter V with different values of D_u where the user QoS requirement is set as $\hat{R}_u = 2$ bps/Hz and the maximum number of matched users of each subchannel is $D_n = 2$. It is seen that, for each value of D_u , the average EE increases and


 Fig. 5. Average EE performance versus simulation time for different values of D_n .

 Fig. 6. Average EE performance versus the parameter V for different values of D_u .

 Fig. 7. The total average transmit capacity versus the parameter V for different values of D_u .

converges to a certain value with increasing V . For the same value of V , a larger value of D_u leads to a larger value of average EE due to the various selection of subchannels.

Fig. 7 illustrates the convergence of the total average capacity with increasing V with different values of D_u . Different from the trend of average EE, for the same value of D_u , the total average transmit capacity decreases with increasing V . For the same value of V , a larger value of D_u results in a larger value of total average transmit capacity. This is because every user has more subchannel selections to guarantee the QoS requirement.

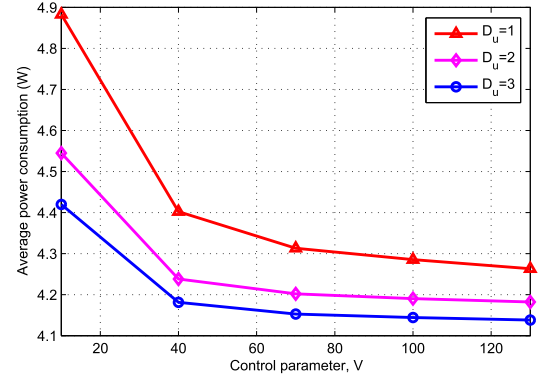
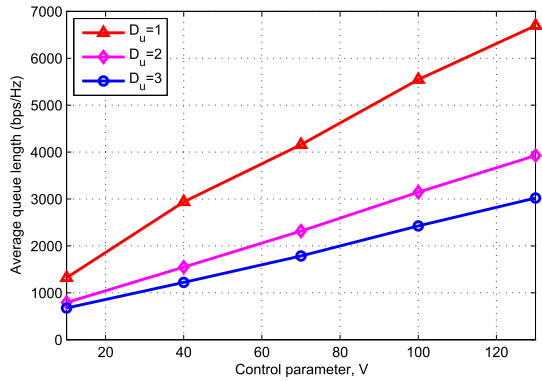

 Fig. 8. The average power consumption versus the parameter V for different values of D_u .

 Fig. 9. The average queue length versus the parameter V for different values of D_u .

Fig. 8 shows the average power consumption versus the parameter V for different values of D_u . The set of constraints is as Fig. 5 and Fig. 6. We see that the average power consumption decreases steadily with increasing V . For the same value of V , different from the trend of the total average capacity, a larger value of D_u results in a lower value of average power consumption. The average power consumption of $D_u = 3$ is 3% smaller than that of $D_u = 1$ when $V = 120$.

Fig. 9 depicts the average queue length versus the parameter V with different values of D_u . The user QoS requirement is set as $\hat{R}_u = 2 \text{ bps/Hz}$ and the maximum number of matched users of each subchannel is $D_n = 2$ in the NOMA scheme. With different values for D_u , the average queue length grows steadily with increasing V . For the same value of V , a larger value of D_u leads to a lower value of average queue length.

In Fig. 10, the average EE versus V with different QoS requirements is shown. We set the maximum number of allocated subchannels of each user as $D_u = 2$ and the maximum number of matched users of each subchannel as $D_n = 2$. As shown in Fig. 10, for the same value of V , a larger value of the QoS requirement results in a smaller converged value of average EE. A smaller value of QoS requirement leads to a larger value of EE due to the fact that a smaller value of QoS requirement enlarges the feasible region of the optimizing variable.

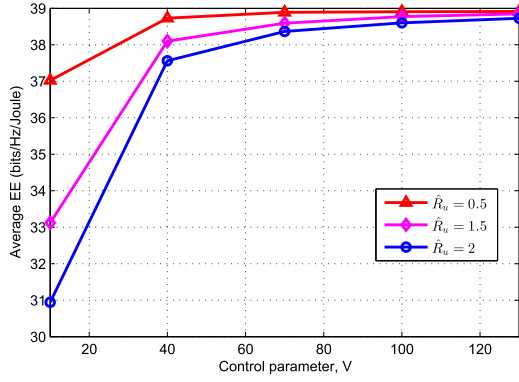


Fig. 10. Average EE performance versus the parameter V for different QoS requirements.

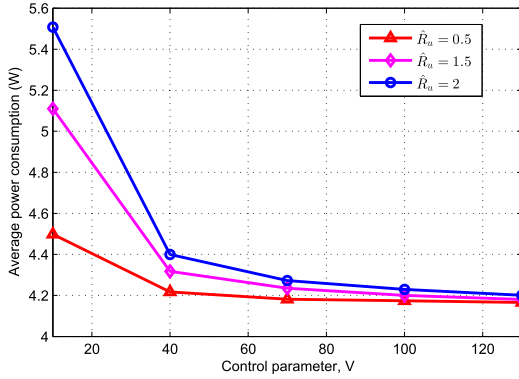


Fig. 11. The average power consumption versus the parameter V for different QoS requirements.

Fig. 11 shows the average power consumption versus V for different QoS requirements. The maximum number of allocated subchannels of each user is $D_u = 2$ and the maximum number of matched users of each subchannel is $D_n = 2$. As the parameter V increases, average power consumption continues to decrease. From Fig. 11, we can observe that for the same value of V , a larger value of the QoS requirement results in a larger value of average power consumption.

In Fig. 12, the total average transmit capacity is evaluated versus V for different QoS requirements. The mean traffic arrival rate is set at $\alpha = 15 \text{ bps/Hz}$ and the other constraints are the same as in Fig. 11. It can be observed that, the trend of this curve is similar to the average power consumption curves in Fig. 11. For the same parameter V , a larger value of the QoS requirement results in a larger value of the total average transmit capacity.

VI. CONCLUSION

We have investigated dynamic resource allocation in downlink NOMA networks. In particular, we have proposed a suboptimal subchannel assignment algorithm based on the two-side matching method. We have formulated the power allocation as a mixed integer programming problem by considering a minimum QoS requirement, and maximum power constraint. Based on the framework of Lyapunov optimization, the problem of energy efficient optimization has been broken

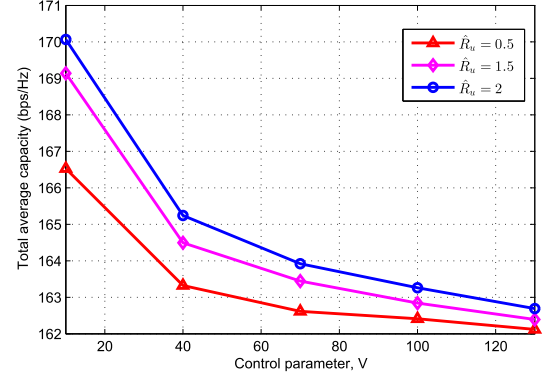


Fig. 12. The total average transmit capacity versus the parameter V for different QoS requirements.

down into three subproblems, two of which are linear and the rest of which can be solved via Lagrangian optimization. The mathematical analysis and simulation results have demonstrated the effectiveness of the proposed algorithms.

APPENDIX A

The proof of Theorem 1 is similar to the method in [27]. For the sake of notational simplicity, we define Θ as the set of of feasible solutions of the optimization problem in (12) and

$\eta_{EE} = \frac{\sum_{k=1}^K \sum_{u=1}^U g_R(\bar{r}_{k,u})}{\bar{p}_{tot}} = \frac{G(\mathcal{P})}{T(\mathcal{P})}$, respectively. Without loss of generality, we define η_{EE}^{opt} and $\{\mathcal{P}^*\} \in \Theta$ as the optimal energy efficiency and the optimal power allocation policy of the original objective function in (12), respectively. We can get the optimal energy efficiency as

$$\begin{aligned} \eta_{EE}^{opt} &= \frac{G(\mathcal{P}^*)}{T(\mathcal{P}^*)} \geq \frac{G(\mathcal{P})}{T(\mathcal{P})}, \forall \{\mathcal{P}\} \in \Theta, \\ &\Rightarrow \begin{cases} G(\mathcal{P}) - \eta_{EE}^{opt} T(\mathcal{P}) \leq 0 \\ G(\mathcal{P}^*) - \eta_{EE}^{opt} T(\mathcal{P}^*) = 0. \end{cases} \end{aligned} \quad (48)$$

Therefore, we can conclude that $\max_{\mathcal{P}} G(\mathcal{P}) - \eta_{EE}^{opt} T(\mathcal{P}) = 0$ is achievable by power allocation policy $\{\mathcal{P}^*\}$ which completes the forward implication.

Then, the converse implication of Theorem 1 is proved as below. Suppose $\mathcal{P}_o^*$ is the optimal power allocation policy of the equivalent objective function such that

$$G(\mathcal{P}_o^*) - \eta_{EE}^{opt} T(\mathcal{P}_o^*) = 0. \quad (49)$$

Then, for any feasible power allocation policy $\{\mathcal{P}\} \in \Theta$, we can obtain the following inequality:

$$G(\mathcal{P}) - \eta_{EE}^{opt} T(\mathcal{P}) \leq G(\mathcal{P}_o^*) - \eta_{EE}^{opt} T(\mathcal{P}_o^*) = 0. \quad (50)$$

The above inequality implies

$$\frac{G(\mathcal{P})}{T(\mathcal{P})} \leq \eta_{EE}^{opt}, \quad \forall \{\mathcal{P}\} \in \Theta \quad (51)$$

and

$$\frac{G(\mathcal{P}_o^*)}{T(\mathcal{P}_o^*)} = \eta_{EE}^{opt}. \quad (52)$$

This implies that the optimal power allocation policy $\{\mathcal{P}_o^*\}$ for the equivalent objective function is also the optimal resource allocation policy for the original objective function.

APPENDIX B

Let Y_1 and Y_2 be the optimal utilities of problems (15) and (17), respectively. The optimal solutions that achieve Y_1 and Y_2 are denoted by X_1 and X_2 , respectively. Since the utility functions of (15) and (17) are non-decreasing concave function which can be expressed as $U(\cdot)$, by Jensen's inequality, we have

$$U(\overline{X_2}) \geq \overline{U(X_2)} = Y_2 \quad (53)$$

Due to the fact that the solution X_2 satisfies the constraint C5, we can write

$$U(\overline{X_1}) \geq U(\overline{X_2}). \quad (54)$$

Moreover, since X_2 also satisfies the constraints of problem (15), we have

$$Y_1 \geq U(\overline{X_1}) \geq Y_2. \quad (55)$$

Since X_1 is an optimal solution to the problem (15), it also satisfies the constraints C1 – C4. By choosing $X_2 = \overline{X_1}$ at each slot, we get

$$Y_2 \geq \overline{U(X_2)} = U(\overline{X_1}) = Y_1. \quad (56)$$

Therefore, $Y_1 = Y_2$ is proved and we can further conclude the equivalence of the problems (15) and (17).

APPENDIX C PROOF OF THEOREM 2

To prove the bounds on the EE and the average queue length, Lemma 1 is introduced below.

Lemma 1: For arbitrary arrival rates, a randomized stationary control policy Π exists and it chooses a feasible control decision independent of current traffic queues and virtual queues. We get the following steady state values:

$$E[r_u^\Pi(t)] = r_u^* \quad (57)$$

$$E[R_u^\Pi(t)] \geq E[r_u^\Pi(t)] + \varepsilon = r_u^* + \varepsilon \quad (58)$$

$$E[\hat{P}_u - p_u^\Pi(t)] \leq \delta \quad (59)$$

$$E[U_{tot}^\Pi(t)] \geq E[p_{tot}^\Pi(t)](\eta_{EE}^{opt} - \delta). \quad (60)$$

Proof: The proof of Lemma 1 is similar to one found in [29].

Substituting (57)–(60) into (22) and taking the limit $\delta \rightarrow 0$, we can get

$$\begin{aligned} \Delta(\Phi(t)) - VE \left\{ \sum_{u \in U} g_R(\gamma_u) - \eta_{EE} p_{tot} \right\} \\ \leq B - V\eta_{EE}^{opt} E(p_{tot}^\Pi(t)) + V\eta_{EE} E(p_{tot}^\Pi(t)) - \varepsilon \sum_{u \in U} Q_u(t). \end{aligned} \quad (61)$$

For $Q_u(t) \geq 0$, the inequality (61) can be further simplified to

$$\begin{aligned} \Delta(\Phi(t)) - VE \left\{ \sum_{u \in U} g_R(\gamma_u) - \eta_{EE} p_{tot} \right\} \\ \leq B - V\eta_{EE}^{opt} E(p_{tot}^\Pi(t)) + V\eta_{EE} E(p_{tot}^\Pi(t)). \end{aligned} \quad (62)$$

Using telescoping sums over $t \in \{0, 1, \dots, T-1\}$ and iterated expectations, we have

$$\begin{aligned} \left\{ E \{L(\Phi(T))\} - E \{L(\Phi(0))\} \right\} \\ \left\{ -VE \{U_{tot} - \eta_{EE}(t)p_{tot}\} \right\} \\ \leq T[B - V\eta_{EE}^{opt} E(p_{tot}^\Pi(t))] + VE(p_{tot}^\Pi(t)) \sum_{t=0}^{T-1} E \{\eta_{EE}\}. \end{aligned} \quad (63)$$

Dividing (63) by VT , we get

$$\begin{aligned} \frac{1}{T} \sum_{t=0}^{T-1} E \{\eta_{EE}(t)p_{tot}\} - \frac{1}{T} \sum_{t=0}^{T-1} E \{U_{tot}\} \\ \leq \frac{B}{V} - \eta_{EE}^{opt} E(p_{tot}^\Pi(t)) \\ + E(p_{tot}^\Pi(t)) \frac{1}{T} \sum_{t=0}^{T-1} E \{\eta_{EE}\} + \frac{E \{L(\Phi(0))\}}{VT}. \end{aligned} \quad (64)$$

Taking the limit $T \rightarrow \infty$, one has

$$\begin{aligned} \lim_{T \rightarrow \infty} \left[\frac{1}{T} \sum_{t=0}^{T-1} E \{\eta_{EE}(t)p_{tot}\} - \frac{1}{T} \sum_{t=0}^{T-1} E \{U_{tot}\} \right] \\ = \bar{U}_{tot} - \bar{U}_{tot} = 0. \end{aligned} \quad (65)$$

We obtain

$$\frac{B}{V} - \eta_{EE}^{opt} E(p_{tot}^\Pi(t)) + \eta_{EE} E(p_{tot}^\Pi(t)) \geq 0. \quad (66)$$

Rearranging (66), we have

$$\eta_{EE} \geq \eta_{EE}^{opt} - \frac{B}{VE(p_{tot}^\Pi(t))} \geq \eta_{EE}^{opt} - \frac{B}{VP_{\min}}. \quad (67)$$

Similarly, taking iterated expectations and applying telescoping sums over $t \in \{0, 1, \dots, T-1\}$ to (61), we get

$$\begin{aligned} \left\{ E \{L(\Phi(T))\} - E \{L(\Phi(0))\} \right\} \\ \left\{ -VE \{U_{tot} - \eta_{EE}(t)p_{tot}\} \right\} \\ \leq T[B - V\eta_{EE}^{opt} E(p_{tot}^\Pi(t))] \\ - \varepsilon \sum_{t=0}^{T-1} \left[\sum_{k \in K} \sum_{u \in U} E \{Q_u(t)\} \right] \\ + VE(p_{tot}^\Pi(t)) \sum_{t=0}^{T-1} E \{\eta_{EE}\}. \end{aligned} \quad (68)$$

Dividing (68) by εT and taking a limit as $T \rightarrow \infty$, we obtain

$$\begin{aligned} \bar{Q} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \left[\sum_{u \in U} E \{Q_u(t)\} \right] \\ \leq \left\{ \frac{B - V\eta_{EE}^{opt} E(p_{tot}^\Pi(t))}{\varepsilon} \right. \\ \left. + \frac{V}{\varepsilon} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E \{\eta_{EE}(t)p_{tot}(t)\} \right\} \\ \leq \frac{B + V(R_{\max} - \eta_{EE}^{opt} P_{\min})}{\varepsilon}. \end{aligned} \quad (69)$$

REFERENCES

- [1] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3618–3631, Oct. 2012.
- [2] D. Yuan, J. Joung, C. K. Ho, and S. Sun, "On tractability aspects of optimal resource allocation in OFDMA systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 863–873, Feb. 2013.
- [3] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE 77th Veh. Technol. Conf.*, Jun. 2013, vol. 53, no. 3, pp. 1–5.
- [4] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537–552, Jan. 2016.
- [5] Z. Yang, J. Cui, X. Lei, Z. Ding, P. Fan, and D. Chen, "Impact of factor graph on average sum rate for uplink sparse code multiple access systems," *IEEE Access*, vol. 4, pp. 6585–6590, 2016.
- [6] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.
- [7] Z. Wei, J. Yuan, D. W. K. Ng, M. ElKashlan, and Z. Ding, "A survey of downlink non-orthogonal multiple access for 5G wireless communication networks," *ZTE Commun.*, vol. 14, no. 4, pp. 17–25, Sep. 2016.
- [8] Z. Ding *et al.*, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.
- [9] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Aug. 2015.
- [10] K. Higuchi and A. Benjebbour, "Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access," *IEICE Trans. Commun.*, vol. 98, no. 3, pp. 403–414, 2015.
- [11] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 768–783, May 2004.
- [12] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.
- [13] M. F. Hanif, Z. Ding, T. Ratnarajah, and G. K. Karagiannis, "A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 76–88, Jan. 2016.
- [14] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1647–1651, Oct. 2015.
- [15] Y. Zhang, H.-M. Wang, T.-X. Zheng, and Q. Yang, "Energy-efficient transmission design in non-orthogonal multiple access," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2852–2857, Mar. 2017.
- [16] 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, Sep. 2016.
- [17] 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.
- [18] M. Al-Imari, P. Xiao, M. A. Imran, and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks," in *Proc. IEEE ISWCS*, Aug. 2014, pp. 781–785.
- [19] B. Wang, L. Dai, Z. Wang, N. Ge, and S. Zhou, "Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2370–2382, Oct. 2017.
- [20] L. P. Qian, Y. Wu, H. Zhou, and X. Shen, "Non-orthogonal multiple access vehicular small cell networks: Architecture and solution," *IEEE Netw.*, vol. 31, no. 4, pp. 15–21, Jul./Aug. 2017.
- [21] J. Zhao, Y. Liu, K. K. Chai, A. Nallanathan, Y. Chen, and Z. Han, "Spectrum allocation and power control for non-orthogonal multiple access in HetNets," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5825–5837, Sep. 2017.
- [22] J. Li, M. Peng, Y. Yu, and Z. Ding, "Energy-efficient joint congestion control and resource optimization in heterogeneous cloud radio access networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9873–9887, Dec. 2016.
- [23] Y. Yu, M. Peng, J. Li, A. Cheng, and C. Wang, "Resource allocation optimization for hybrid access mode in heterogeneous networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 1243–1248.
- [24] Q. Chen, G. Yu, R. Yin, A. Maaref, G. Y. Li, and A. Huang, "Energy efficiency optimization in licensed-assisted access," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 723–734, Apr. 2016.
- [25] H. Zhang, B. Wang, C. Jiang, K. Long, A. Nallanathan, and V. C. M. Leung, "Energy efficient dynamic resource allocation in NOMA networks," in *Proc. IEEE Global Commun. Conf.*, Singapore, Dec. 2017, pp. 1–5.
- [26] L. Lei, D. Yuan, C. K. Ho, and S. Sun, "Power and channel allocation for non-orthogonal multiple access in 5G systems: Tractability and computation," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8580–8594, Dec. 2016.
- [27] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967. [Online]. Available: <http://www.jstor.org/stable/2627691>
- [28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [29] Y. Li, M. Sheng, Y. Shi, X. Ma, and W. Jiao, "Energy efficiency and delay tradeoff for time-varying and interference-free wireless networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 5921–5931, Nov. 2014.



Haijun Zhang (M'13–SM'17) was a Post-Doctoral Research Fellow with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver Campus, Canada. He is currently a Full Professor with the University of Science and Technology Beijing, China. He received the IEEE ComSoc Young Author Best Paper Award in 2017. He serves/served as the General Co-Chair for GameNets 2016, the GLOBECOM 2017 Workshop on LTE-U, and the ICC 2018 (ICC 2017, GLOBECOM 2017) Workshop on 5G Ultra Dense Networks, the TPC Co-Chair for INFOCOM 2018 Workshop on IECCON, and the Symposium Chair for GLOBECOM 2019. He serves as the Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE 5G Tech Focus, and EURASIP Journal on Wireless Communications, and serves/served as a Leading Guest Editor for the IEEE Communications Magazine and the IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING.



Baobao Wang received the B.S. degree in communication engineering from the Beijing University of Chemical Technology, Beijing, China, in 2015, where she is currently pursuing the M.S. degree. Her research interests include resource allocation, power control, and energy efficiency in wireless communication.



Chunxiao Jiang (S'09–M'13–SM'15) received the B.S. degree (Hons.) in information engineering from Beihang University in 2008 and the Ph.D. degree (Hons.) in electronic engineering from Tsinghua University in 2013. From 2013 to 2016, he held a post-doctoral position with the Department of Electronic Engineering, Tsinghua University, during which he visited the University of Maryland at College Park and the University of Southampton. He is currently an Assistant Professor with the Tsinghua Space Center, Tsinghua University. He was a recipient of the IEEE GLOBECOM Best Paper Award in 2013, the IEEE GlobalSIP Best Student Paper Award in 2015, the IEEE IWCMC Best Paper Award in 2017, and the IEEE Communications Society Young Author Best Paper Award in 2017.



Keping Long (SM'06) received the M.S. and Ph.D. degrees from the University of Electronic Science and Technology of China, Chengdu, in 1995 and 1998, respectively.

From 1998 to 2000, he was a Post-Doctoral Research Fellow with the National Laboratory of Switching Technology and Telecommunication Networks, Beijing University of Posts and Telecommunications (BUPT), China. From 2000 to 2001, he was an Associate Professor at BUPT. From 2001 to 2002, he was a Research Fellow with the

ARC Special Research Centre for Ultra Broadband Information Networks, The University of Melbourne, Australia. He is currently a Professor and the Dean with the School of Computer and Communication Engineering, University of Science and Technology Beijing. He has authored over 200 papers, 20 keynote speeches, and invited talks at international and local conferences. His research interests are optical Internet technology, new generation network technology, wireless information networks, value-added services, and secure technology of networks. He has been a TPC or an ISC Member of COIN 2003–2010, the IEEE IWCN 2010, ICON in 2004 and 2006, respectively, and APOC in 2004, 2006, and 2008, respectively. He is a member of the Editorial Committees of Sciences in China Series F and China Communications. He has been the Co-Chair of the Organization Committee for IWCMC 2006, the TPC Chair of COIN in 2005 and 2008, and the TPC Co-Chair of COIN in 2008 and 2010. He received the National Science Fund for Distinguished Young Scholars of China in 2007 and selected as the Chang Jiang Scholars Program Professor of China in 2008.



Arumugam Nallanathan (S'97–M'00–SM'05–F'17) was an Assistant Professor with the Department of Electrical and Computer Engineering, National University of Singapore, from 2000 to 2007. He was with the Department of Informatics, Kings College London, from 2007 to 2017, where he was a Professor of wireless communications from 2013 to 2017 and has been a Visiting Professor since 2017. He has been a Professor of wireless communications and the Head of the Communication Systems Research Group, School of Electronic

Engineering and Computer Science, Queen Mary University of London, since 2017. He has published nearly 400 technical papers in scientific journals and international conferences. His research interests include 5G wireless networks, Internet of Things, and molecular communications. He was a co-recipient of the Best Paper Awards presented at the IEEE International Conference on Communications 2016 and the IEEE Global Communications Conference 2017. He is an IEEE Distinguished Lecturer. He has been selected as a Web of Science Highly Cited Researcher in 2016.

He received the IEEE Communications Society SPCE Outstanding Service Award in 2012 and the IEEE Communications Society RCC Outstanding Service Award in 2014. He served as the Chair for the Signal Processing and Communication Electronics Technical Committee of the IEEE Communications Society and the technical program chair and a member of the technical program committees in numerous IEEE conferences. He was an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2006 to 2011, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2006 to 2017, the IEEE WIRELESS COMMUNICATIONS LETTERS, and the IEEE SIGNAL PROCESSING LETTERS. He is an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS.



Victor C. M. Leung (S'75–M'89–SM'97–F'03) received the B.A.Sc. (Hons.) and Ph.D. degrees in electrical engineering from The University of British Columbia (UBC) in 1977 and 1982, respectively. He received the APEBC Gold Medal as the head of the 1977 graduating class in the Faculty of Applied Science. He received a Canadian Natural Sciences and Engineering Research Council Post-graduate Scholarship to support his graduate studies.

From 1981 to 1987, he was a Senior Member of Technical Staff and a Satellite System Specialist at MPR Teltech Ltd., Canada. In 1988, he was a Lecturer with the Department of Electronics, The Chinese University of Hong Kong. He returned to UBC as a Faculty Member in 1989, where he is currently a Professor and the TELUS Mobility Research Chair in Advanced Telecommunications Engineering with the Department of Electrical and Computer Engineering. He has co-authored over 1100 journal/conference papers, 40 book chapters, and co-edited 14 book titles. Several of his papers had been selected for best paper awards. His research interests are in the broad areas of wireless networks and mobile systems.

Dr. Leung is a fellow of the Royal Society of Canada, the Engineering Institute of Canada, and the Canadian Academy of Engineering. He received the IEEE Vancouver Section Centennial Award, the 2011 UBC Killam Research Prize, the 2017 Canadian Award for Telecommunications Research, and the 2018 IEEE ComSoc TGCC Distinguished Technical Achievement Recognition Award. He co-authored papers that received the 2017 IEEE ComSoc Fred W. Ellersick Prize, the 2017 IEEE Systems Journal Best Paper Award, and the 2018 IEEE ComSoc CSIM Best Journal Paper Award. He has served on the Editorial Boards of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS—Wireless Communications Series and Series on Green Communications and Networking, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON COMPUTERS, the IEEE WIRELESS COMMUNICATIONS LETTERS, and the *Journal of Communications and Networks*. He is serving on the Editorial Boards of the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, the IEEE TRANSACTIONS ON CLOUD COMPUTING, the IEEE ACCESS, *Computer Communications*, and several other journals. He has guest edited many journal special issues and provided leadership to the organizing committees and technical program committees of numerous conferences and workshops. He was a Distinguished Lecturer of the IEEE Communications Society. He is a registered Professional Engineer in the Province of British Columbia, Canada.



H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 to 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty of Princeton University, where he is currently the Michael Henry Strater University Professor of Electrical Engineering. From 2006 to 2016, he served as the Dean of Princeton's School of Engineering and Applied Science. He has also held visiting appointments at several other universities,

including most recently at Berkeley and Cambridge. His research interests are in the areas of information theory and signal processing, and their applications in wireless networks, energy systems, and related fields. Among his publications in these areas is the recent book *Information Theoretic Security and Privacy of Information Systems* (Cambridge University Press, 2017).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and a foreign member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. He received the Marconi and Armstrong Awards of the IEEE Communications Society in 2007 and 2009, respectively. Recent recognition of his work includes the 2017 IEEE Alexander Graham Bell Medal, Honorary Professorships at Peking University and Tsinghua University, both conferred in 2017, and a D.Sc. *honoris causa* from Syracuse University also awarded in 2017.