Hybrid NOMA Offloading in Multi-User MEC Networks

Zhiguo Ding, *Fellow, IEEE*, Dongfang Xu, *Graduate Student Member, IEEE*, Robert Schober, *Fellow, IEEE*, and H. Vincent Poor, *Life Fellow, IEEE*

***Abstract*— Non-orthogonal multiple access (NOMA) assisted mobile edge computing (MEC) has recently attracted significant attention due to its superior capability to reduce the energy consumption and the latency of MEC offloading. In this paper, a general hybrid NOMA-MEC offloading strategy is proposed, which includes conventional orthogonal multiple access (OMA) and pure NOMA based offloading as special cases. A multi-objective optimization problem is formulated to minimize the energy consumption for MEC offloading, and a low-complexity resource allocation solution is derived and shown to be Pareto-optimal. Furthermore, by analyzing the properties of the obtained resource allocation solution, important insights regarding NOMA-MEC offloading are obtained. For example, it is proved that pure NOMA-MEC offloading cannot outperform hybrid NOMA-MEC. In addition, a precise condition under which NOMA-MEC outperforms OMA-MEC is established, and shown to match the one previously developed for the two-user special case. Furthermore, the developed analytical results also establish an interesting analogy between the proposed hybrid NOMA-MEC power allocation scheme and the well-known water-filling strategy.**

***Index Terms*— Non-orthogonal multiple access (NOMA), mobile edge computing (MEC), multi-objective optimization, Pareto optimality.**

I. Introduction

MOBILE edge computing (MEC) has been recognized as an important enabling technique for the next generation of wireless networks [1]–[5]. The key idea behind MEC is to employ the infrastructure close to the edge of mobile networks, such as base stations, as computing servers, such that the mobile users can offload their computationally-intensive delay-critical tasks to these servers. Because of the superior computing capabilities of the MEC servers, these tasks can be computed faster than when they are computed locally. From the energy consumption perspective, MEC offers mobile users the benefit of prolonging their battery lifetimes, since they do not have to spend energy for local computing. The concept of MEC is particularly important for the application of machine learning in the context of the wireless Internet of Things (IoT), where devices often are energy constrained and have limited computing capabilities [6]–[8]. Instead of relying on the IoT devices to locally compute machine learning tasks, MEC offers an energy efficient and low-latency alternative for task computation.

Energy and spectrally efficient offloading is key to the success of MEC networks, which motivates the application of non-orthogonal multiple access (NOMA) to MEC offloading [9], [10]. In particular, recall that with conventional orthogonal multiple access (OMA), only a single user can be served in each resource block, such as a time slot or a frequency channel, whereas the use of NOMA can ensure that multiple users are served simultaneously in a single resource block. As a result, NOMA-MEC is more spectrally efficient than OMA-MEC, since a single U resource block can be used to support multiple users' offloading. Furthermore, both the delay and the energy consumption of MEC offloading can also be significantly reduced by applying NOMA, as shown in [11]. In [12], the delay minimization problem for NOMA-MEC offloading was investigated for a single-cell multi-user network, which was then generalized to a multi-cell multi-user network in [13]. In [14]–[16], NOMA-MEC networks employing orthogonal frequency division multiplexing (OFDM) were studied, where sophisticated approaches for joint power and subcarrier allocation were developed. In [17]–[19], the use of unmanned aerial vehicles (UAVs) in NOMA-MEC networks was considered, and various designs for joint MEC resource allocation and UAV trajectory planning were developed. In [20] and [21], security provisioning for NOMA-MEC networks was investigated, where the presence of passive eavesdroppers was assumed and different approaches for security enhancement compared to OMA-MEC were developed. To facilitate energy-constrained IoT, the application of wireless power transfer was considered in [22] and [23], where sophisticated algorithms for joint task allocation and delay minimization were proposed.

In the aforementioned works, pure NOMA-MEC offloading was used, i.e., all users had to start and finish their offloading

Manuscript received September 17, 2021; revised December 18, 2021; accepted December 27, 2021. Date of publication January 12, 2022; date of current version July 12, 2022. The work of Zhiguo Ding was supported by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/P009719/2, in part by the H2020-MSCA-RISE-2020 under Grant 101006411, and in part by the Friedrich Wilhelm Bessel Research Prize from the Alexander von Humboldt Foundation. The work of Robert Schober was supported by the BMBF Project 6G-RIC under Grant 16KISK023. The work of H. Vincent Poor was supported by the U.S. National Science Foundation under Grant CCF-1908308. The associate editor coordinating the review of this article and approving it for publication was N. Kato. *(Corresponding author: Zhiguo Ding.)*

Zhiguo Ding is with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544 USA, and also with the School of Electrical and Electronic Engineering, The University of Manchester, Manchester M13 9PL, U.K. (e-mail: zhiguo.ding@manchester.ac.uk).

Dongfang Xu and Robert Schober are with the Institute for Digital Communications, Friedrich-Alexander-University Erlangen-Nurnberg (FAU), 91054 Erlangen, Germany (e-mail: dongfang.xu@fau.de; robert.schober@fau.de).

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

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2021.3139932.

Digital Object Identifier 10.1109/TWC.2021.3139932

at the same time. Unlike these existing works, this paper considers a general hybrid NOMA scheme for MEC offloading, where not all users start and finish their offloading at the same time, i.e., it is possible for a user to first offload parts of its task together with other users, i.e, in the pure NOMA mode, and then offload the remainder of its task by solely occupying a resource block, i.e., in the OMA mode. The use of hybrid NOMA-MEC offloading was investigated for the special case of two users in [24]–[26], where several insightful properties of MEC offloading were revealed. For example, the authors of [24], [25] showed that the use of hybrid NOMA offloading always yields an energy consumption that does not exceed that of pure NOMA offloading, where all the users have to finish offloading concurrently. Furthermore, the authors of [24] provided a precise condition under which NOMA-MEC outperforms OMA-MEC. In particular, if one user's task deadline is less than two times the other user's deadline, NOMA-MEC can outperform OMA-MEC. This condition is intuitive as explained in the following. If a user's deadline is not demanding, the user can afford to wait until the other user finishes offloading, and then use OMA for its own offloading, which avoids transmission in the presence of the strong interference caused by the other user. However, these conclusions were developed for the special case of two users, and it is not clear whether they are applicable for more general multi-user networks, which provides the motivation for this paper.

In particular, this paper focuses on a general multi-user MEC network, and the contributions of this paper are three-fold, as listed in the following:

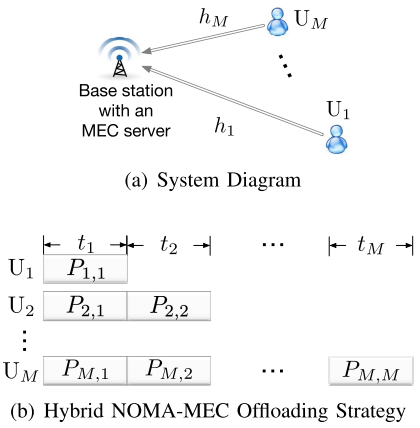
- A general multi-user hybrid NOMA-MEC offloading strategy is first developed, for which the two-user hybrid offloading strategy designed in [24], [25], [27] and the iterative two-user NOMA scheme developed in [26] can be viewed as special cases. In particular, for the proposed NOMA-MEC strategy, the users' offloading is scheduled according to the urgency of their tasks. Unlike for OMA-MEC, when a user with a more urgent task offloads, other users with less urgent tasks can still carry out offloading. Furthermore, unlike pure NOMA-MEC, which forces the users to start and finish their offloading simultaneously, the proposed strategy offers the users the opportunity to offload their tasks more flexibly, e.g., a user can wait until all the other users with more urgent tasks have finished their offloading. As a result, both pure NOMA-MEC and OMA-MEC are special cases of the proposed general hybrid offloading strategy.
- A multi-objective optimization problem is formulated in order to minimize the users' energy consumption for MEC offloading while meeting delay and transmit power constraints. We note that it is challenging to solve the formulated optimization problem because of its multi-objective and non-convex nature. Nevertheless, a low-complexity successive optimization algorithm is proposed by exploiting an important feature of the considered NOMA-MEC network, namely that the transmission strategy of a user with a less demanding task deadline has no impact on the data rate of a user with a more urgent task. By applying the proposed successive optimization algorithm, a closed-form expressions for the users' offloading power and time allocation parameters can be obtained.
- The properties of the proposed successive optimization algorithm are analyzed, which provides greater insight into NOMA-MEC. For example, the developed analytical results show that pure NOMA cannot outperform hybrid NOMA, which is consistent with the insight obtained for the two-user special case in [24], [25]. A precise condition for NOMA-MEC to outperform OMA-MEC is also established, which reveals that the number of bits to be offloaded and the existing co-channel interference play important roles when deciding which of the two modes, namely OMA-MEC and NOMA-MEC, is preferable. This result is different from the conclusion reached for the two-user case in [24], [25]. However, it is worth noting that this new condition can be degraded to the one developed for the two-user case [24]. Another important insight obtained in this paper is that each user adopts a power allocation strategy similar to the well-known water-filling strategy [28]. Furthermore, it is proved that the proposed successive optimization algorithm realizes one of the Pareto optimal solutions for the considered multi-objective optimization problem, and the provided simulation results confirm the Pareto-optimality of the obtained solution.

The remainder of this paper is organized as follows. In Section II, the system model for the considered NOMA-MEC system is described and a multi-objective optimization problem for the energy consumption of MEC offloading is formulated. In Section III, the proposed successive optimization algorithm is described and its properties are analyzed in order to obtain an insightful understanding of NOMA-MEC. Computer simulations are presented in Section IV, and the paper is concluded in Section V. Finally, all the proofs are collected in the appendix.

II. System Model

Consider a multi-user MEC network with M users, denoted by U_m, $1 \leq m \leq M$, and a base station equipped with an MEC server, as shown in Fig. 1(a). Each node is assumed to be equipped with a single antenna. Each user needs to compute a computationally-intensive and latency-critical task which is to be offloaded to the base station. Denote U_m's deadline by D_m. Without loss of generality, we assume that the users are ordered according to the urgency of their task deadlines, i.e., $D_1 \leq \cdots \leq D_M$.

A general hybrid NOMA-MEC offloading strategy is proposed in this paper, where the users are scheduled to complete their offloading according to the urgency of their tasks, as shown in Fig. 1(b). In particular, during the first t_1 seconds, U_1 is asked to complete its offloading, because its deadline is the most demanding one. In addition to U_1, the other users are also allowed to carry out offloading during the first t_1 seconds. During the next t_2 seconds, U_2 is asked to complete



(a) System Diagram

(b) Hybrid NOMA-MEC Offloading Strategy

Fig. 1. Illustration of the considered NOMA-MEC system.

its offloading, where the other users, U_m, $2 \leq m \leq M$, can continue offloading their tasks simultaneously. This offloading strategy continues in a successive manner according to the users' delay requirements, as shown in Fig. 1(b). During the last t_M seconds, only U_M is served, because all other users should have already completed their offloading by then. Denote the users' transmit powers during t_n by $P_{m,n}$, $1 \leq n \leq m$, $1 \leq m \leq M$.

We note that both pure OMA and pure NOMA offloading are special cases of this hybrid NOMA offloading scheme. For example, by setting $t_m = 0$ for $m \geq 2$ (or $P_{m,n} = 0$ for $n \geq 2$ and $m \geq 2$), all users need to complete their offloading within the first t_1 seconds, which corresponds to a pure NOMA scheme. If $P_{m,n} = 0$ for $n < m$, i.e., only a single user is allowed to transmit at a time, the hybrid NOMA scheme is degraded to the pure OMA scheme. As shown in [24], for OMA offloading, for minimization of the offloading energy consumption, the optimal delay is given by $t_m^{\mathrm{OMA}} = D_m - D_{m-1}$, and the optimal power allocation coefficient, denoted by P_m^{OMA}, is obtained by ensuring $(D_m - D_{m-1}) \log\left(1 + P_m^{\mathrm{OMA}} |h_m|^2\right) = N$, where the users' channel coefficients are denoted by h_m and it is assumed that all users' tasks contain the same number of nats, denoted by N.[1] Impairment by additive white Gaussian noise with normalized variance is assumed at all receivers.

By applying the proposed hybrid NOMA-MEC scheme, at t_n, the base station receives the following signal:

$$y_n = \sum_{m=n}^{M} \sqrt{P_{m,n}} h_m s_{m,n} + w_n, \tag{1}$$

where $s_{m,n}$ denotes the signal sent by U_m during t_n, and w_n denotes the white Gaussian noise. In this paper, we assume that the successive interference cancellation (SIC) decoding order is determined by the users' task deadlines. In particular, at t_n, the base station first decodes U_m's signal, when $m > n$, i.e., the signal from the user with the less urgent task, before decoding U_{m-1}'s signal, i.e., the signal from the user with the more urgent task. By using this SIC decoding order, it is guaranteed that U_n experiences the same performance as with OMA at t_n. As a result, at t_n, U_m's signal is decoded in the $(M - m + 1)$-th SIC step with the following offloading data rate:

$$R_{m,n} = \log\left(1 + \frac{P_{m,n}|h_m|^2}{\sum_{j=n}^{m-1} P_{j,n}|h_j|^2 + 1}\right). \tag{2}$$

As a consequence, at t_n, U_n's signal is decoded last with the offloading data rate of $R_{n,n} = \log\left(1 + P_{n,n}|h_n|^2\right)$, which means that U_n experiences interference-free transmission as in OMA.

MEC offloading imposes two constraints. One is to ensure that U_m can offload all its N nats by the end of t_m, i.e.,

$$\sum_{n=1}^{m} t_n R_{m,n} \geq N, \tag{3}$$

and the other constraint is to meet the deadline of MEC offloading, i.e.,

$$\sum_{n=1}^{m} t_n \leq D_m. \tag{4}$$

Define $E_m \triangleq \sum_{n=1}^{m} P_{m,n} t_n$, which denotes U_m's overall energy consumption. Furthermore, define $\mathbf{p}_m = \left[p_{m,1} \cdots p_{m,m}\right]^T$, which collects U_m's power allocation coefficients, and $\mathbf{x} \triangleq \left[t_2 \cdots t_M \ \mathbf{p}_2^T \cdots \mathbf{p}_M^T\right]^T$, which collects the M users' time and power allocation coefficients, where $\mathbf{A}^T$ denotes the transpose of $\mathbf{A}$. In this paper, the following multi-objective energy minimization problem is considered:

$$\min_{\mathbf{x}} \quad \mathbf{E}_M \triangleq \left[E_2 \cdots E_M\right]^T \tag{P1a}$$

$$\text{s.t.} \quad \sum_{n=1}^{m} t_n R_m, n \geq N, \quad 2 \leq m \leq M \tag{P1b}$$

$$\sum_{n=1}^{m} t_n \leq D_m, \quad t_m \geq 0, \quad 2 \leq m \leq M \tag{P1c}$$

$$0 \leq P_{m,n} \leq P_m^{\mathrm{OMA}}, \quad 2 \leq m \leq M, 1 \leq n \leq m, \tag{P1d}$$

where P_m^{OMA}, a user's transmit power in OMA, is used in (P1d) in order to facilitate the performance comparison between OMA-MEC and NOMA-MEC. Another benefit to use P_m^{OMA} as the transmit power budget is that the optimal choices for U_1's transmit parameters, t_1 and $P_{1,1}$, can be straightforwardly obtained as follows: $t_1^* = D_1$ and $P_{1,1}^* = P_1^{\mathrm{OMA}}$, which is the reason why only U_m's parameters, $m > 1$, are optimized in problem P1.

Remark 1: Problem P1 is a typical multi-objective optimization problem, where there are potential conflicts between the multiple users' objectives. For example, one user's choices for its power allocation coefficients may reduce its own offloading energy consumption, but could increase other users' energy consumption. For a multi-objective optimization problem, it is of interest to find Pareto optimal solutions which are defined as follows. Define $\mathcal{E}_M$ as a set collecting all feasible $\mathbf{E}_M$. Denote a feasible solution by $\bar{\mathbf{x}}$, and the corresponding objective value

[1] In this paper, a homogenous scenario is assumed, where each user's task is assumed to contain N nats. In practice, the users' tasks need to be divided into packets for the preparation of uplink transmission. For all packets sent in a communication system, it is reasonable to assume that their sizes are identical, even if they originate from different users.

by $\bar{\mathbf{E}}_M$. If $\bar{\mathbf{x}}$ is Pareto optimal, there is no other element in $\mathcal{E}_M$ that dominates $\bar{\mathbf{E}}_M$, i.e., there is no $\mathbf{z} \in \mathcal{E}_M$ and $\mathbf{z} \neq \bar{\mathbf{E}}_M$ satsifying $\mathbf{z} \prec \bar{\mathbf{E}}_M$, where $\prec$ denotes an element-wise inequality. In the other words, $\bar{\mathbf{E}}_M$ is a minimal element of $\mathcal{E}_M$ [29].

Remark 2: One approach to find the Pareto optimal solutions of problem P1 is to apply the scalarization method and convert the multi-objective optimization problem to a single-objective optimization problem, e.g., using a weighted sum, $\sum_{m=1}^{M} w_m E_m$, as the objective function, where w_m denotes the weights [29]. Different Pareto-optimal solutions can be obtained by varying the weights, w_m. We note that if scalarization is employed, problem P1 is still challenging to solve due to the following two difficulties. First, $P_{m,n}$ and t_m are coupled in the objective function as well as in the constraints. Second, $R_{m,n}$ in (P1b) is not a concave function of the power allocation coefficients. To avoid these obstacles, in the following section, a low-complexity algorithm for solving problem P1 will be introduced, and its Pareto-optimality will be analyzed.

III. A Low-Complexity Successive Resource Allocation Algorithm

In this section, first, a low-complexity successive algorithm is proposed to solve problem P1, and then the properties of the solution obtained with the algorithm are revealed.

A. Description of Successive Resource Allocation

An important observation for the considered NOMA-MEC problem is that due to the use of SIC, U_n's choices for t_n and $P_{n,m}$ have no impact on U_m's data rate, $m < n$. An extreme example is U_1's data rate which is $R_{1,1} = t_1 \log(1+P_{1,1}|h_1|^2)$ and depends on t_1 and $P_{1,1}$ only. This motivates the use of a successive optimization strategy, where U_m's transmission parameters are optimzied after U_{m-1}'s. As a result, E_m can be minimized by focusing on optimizing U_m's parameters only, because U_n's parameters, $1 \leq n \leq m-1$, have already been optimized and U_i's parameters, $i > m$, have no impact on U_m's data rate.

Based on the above discussions, the proposed low-complexity algorithm decomposes problem P1 into a series of subproblems and solves the subproblems in a successive manner, as shown in Algorithm 1, where $\mathbf{p}_m^* = \left[p_{m,1}^* \cdots p_{m,m}^*\right]^T$, and problem P2 is defined as follows:

$$\min_{\mathbf{p}_m, t_m} \quad E_m \tag{P2a}$$

$$\text{s.t.} \quad \sum_{n=1}^{m} t_n R_{m,n} \geq N \tag{P2b}$$

$$0 \leq t_m \leq D_m - \sum_{n=1}^{m-1} t_n \tag{P2c}$$

$$0 \leq P_{m,n} \leq P_m^{\mathrm{OMA}}, \quad 2 \leq n \leq m. \tag{P2d}$$

Algorithm 1 can be viewed as a greedy approach which decomposes problem P1 into the series of the subproblems shown in (P2) and then solves these subproblems in a successive manner. An advantage of Algorithm 1 is that it can be executed with low complexity.[2] Intuitively, the solution obtained with Algorithm 1 is expected to be a suboptimal solution of problem P1. However, surprisingly Algorithm 1 realizes one of the Pareto optimal solutions of problem P1, as shown at the end of this section.

Algorithm 1 Low-Complexity Successive Optimization

1: Set $P_{1,1}^* = P_1^{\mathrm{OMA}}$, $t_1^* = D_1$
2: $m = 1$
3: **while** $m < M$ **do**
4: $\quad m = m + 1$.
5: $\quad$ Find the optimal solutions for t_m^*, $P_{m,n}^*$, $1 \leq n \leq m$, by solving problem P2
6: **end**
7: The outcome of the algorithm is given by $\mathbf{x}^* \triangleq \left[t_2^* \cdots t_M^* \ \mathbf{p}_2^{*T} \cdots \mathbf{p}_M^{*T}\right]^T$

As discussed before, the successive nature of Algorithm 1 comes from the fact that problem P2 is optimized by assuming that U_n's parameters, $1 \leq n \leq m$, have already been optimized. This implies that t_n and $P_{n,i}$, $1 \leq n < m$, shown in problem P2 are constant and fixed.[3] We note that Algorithm 1 is initialized with the choices, $P_{1,1}^* = P_1^{\mathrm{OMA}}$, $t_1^* = D_1$, because P_1^{OMA} and D_1 are the optimal solutions for U_1's transmission parameters.

We note that problem P2 is not convex due to the fact that t_m and $P_{m,n}$ are coupled in the objective and constraint functions. In order to reduce the number of the Lagrange multipliers to be used in the Karush–Kuhn–Tucker (KKT) conditions, the upper bound on $P_{m,n}$ is first removed, which means that the optimization problem considered in (P2) can be recast as follows:

$$\min_{\mathbf{p}_m, t_m} \quad P_{m,m} t_m + \sum_{n=1}^{m-1} P_{m,n} t_n \tag{P3a}$$

$$\text{s.t. (P2b), (P2c)}$$

$$0 \leq P_{m,n}, 1 \leq n \leq m. \tag{P3b}$$

We note that problem P3 is still equivalent to problem P2, if $D_m - D_{m-1} \leq -\frac{N}{\log a_{m,1}}$, $2 \leq m \leq M$, as will be shown in Lemma 5.

We further note that for the first $m-1$ time slots, i.e., t_i, $1 \leq i \leq m-1$, it is possible that U_m does not want to use

[2]Assuming that the base station knows each users' channel state information (CSI) and task deadline, the base station only needs to carry out $M-1$ steps to implement Algorithm 1, where each step is to find the optimal solution of subproblem P2 and the associated computational complexity is moderate due to the fact that subproblem P2 is convex. For example, as shown in Corollary 2, for the case of $D_m - D_{m-1} \leq -\frac{N}{\log a_{m,1}}$, $2 \leq m \leq M$, the optimal solution for subproblem P2 can be obtained with closed-form expressions, i.e., $P_{m,m}^* = \frac{c_m - 1}{|h_m|^2}$ and $P_{m,n}^* = \frac{a_{m,n} c_m - 1}{a_{m,1}|h_m|^2}$ for $n < m$, where $c_m = e^{\frac{N - D_{m-1} \log a_{m,1}}{D_m}}$, and $a_{m,n} = \frac{1}{\sum_{j=n}^{m-1} P_{j,n}|h_j|^2 + 1}$, which means that the overall computational complexity is determined by the complexity to compute $(M-1)$ exponential functions (i.e., $c_m, 2 \leq m \leq M$) and $3(M-1)$ additional multiplications to generate $P_{m,m}^*$ and $P_{m,n}^*$, $2 \leq m \leq M$.

[3]More rigorously, the notations, t_n^* and $P_{n,i}^*$, $1 \leq n < m$, should be used in problem P2, but the supscript, $*$, is omitted for notational simplicity.

all of them for offloading. Define $\mathcal{S}_m$ as the set collecting the indices of the time slots, in which U_m decides to transmit, e.g., $\mathcal{S}_m = \{1, \cdots, m-1\}$ means that U_m transmits continuously during the first $m-1$ time slots, and $\mathcal{S}_m = \emptyset$ means that U_m relies on t_m only.[4] Define also $a_{m,n} = \frac{1}{\sum_{j=n}^{m-1} P_{j,n}|h_j|^2+1}$. Then, a closed-form solution of problem P3 is provided in the following lemma.

Lemma 1: For the optimization problem in (P3), *if* $\min\{a_{m,1}, \cdots, a_{m,m-1}\} \geq e^{-\frac{N-\sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k}}$, *the optimal solution for the power allocation is one of the following two solutions. The first solution is based on pure OMA, i.e.,* $P_{m,n}^{O*} = 0$, *for* $1 \leq n \leq m-1$, *and* $P_{m,m}^{O*} = \frac{e^{\frac{N}{t_m}}-1}{|h_m|^2}$. *The second solution is based hybrid NOMA as shown in the following:*

$$P_{m,n}^{*} = \begin{cases} \dfrac{e^{\frac{N-\sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k}} - 1}{|h_m|^2}, & \text{if } n = m \\ \dfrac{a_{m,n} e^{\frac{N-\sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k}} - 1}{a_{m,n}|h_m|^2}, & \text{if } n < m \\ & \& \; n \in \mathcal{S}_m \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Otherwise, the pure OMA power allocation solution is optimal. For both OMA and hybrid NOMA, the optimal choice of t_m *is the same and given by* $t_m^* = D_m - D_{m-1}$.

Proof: See Appendix A. □

Remark 3: The implementation of Lemma 1 requires the *a priori* knowledge of $\mathcal{S}_m$. Therefore, a straightforward way to use Lemma 1 is to find the potential choices of $\mathcal{S}_m$ and compare their corresponding energy consumption by using the lemma. There is a more computationally efficient alternative, as explained in the following. Lemma 1 shows that the optimal solution for t_m is always $D_m - D_{m-1}$ and does not depend on $P_{m,n}$. By substituting t_m^* into problem P3, problem P3 becomes a convex problem with regard to $P_{m,n}$, and hence off-the-shelf convex optimization solvers can be straightforwardly used to find $P_{m,n}$ without *a priori* knowledge of $\mathcal{S}_m$.

Remark 4: Recall that the considered offloading scheme provides a general framework, and includes three possible power allocation strategies, namely pure NOMA, pure OMA, and hybrid NOMA. The proof of Lemma 1 shows that the pure NOMA strategy cannot outperform hybrid NOMA, and hence can be ignored. This observation is consistent with the results obtained for the two-user special case [24], [25].

Remark 5: While Lemma 1 provides closed-form expressions for the power allocation solutions for pure OMA and hybrid NOMA, it is not clear which of the two yields a better performance. For the two-user special case, it was shown in [24] that hybrid NOMA outperforms OMA if one user's task deadline is less than twice the other user's deadline. A similar condition for a more general setting will be established in the next subsection.

Remark 6: For the two-user case considered in [24], [25], the hybrid NOMA strategy is simple, i.e., one user chooses to transmit in one time slot and the other user transmits in two time slots. The hybrid NOMA strategy provided in Lemma 1 is more complicated. For example, U_M might choose to transmit at t_1 and t_M, and remain silent during the other time slots, i.e., $\mathcal{S}_M = \{1, M\}$. In the next subsection, Lemma 1 will be further simplified by proving that among all the hybrid NOMA strategies with different $\mathcal{S}_m$, the continuous transmission strategy, where $\mathcal{S}_m = \{1, \cdots, m-1\}$, results in the lowest energy consumption.

B. Properties of Successive Resource Allocation

In this section, the properties of successive resource allocation are analyzed by focusing on the important case when $D_n - D_{n-1} \leq -\frac{N}{\log a_{n,1}}$, for $2 \leq n \leq M$, which is the feasibility condition for the hybrid NOMA solution with $\mathcal{S}_m = \{1, \cdots, m-1\}$ as will be explained later. One reason to focus on this particular case is that it corresponds to the important time-critical situation, i.e., the duration left for U_n's offloading in OMA, $D_n - D_{n-1}$, is small. The other reason is that more concise conclusions about NOMA-MEC offloading can be obtained.

Recall that $a_{m,n} = \frac{1}{\sum_{j=n}^{m-1} P_{j,n}|h_j|^2+1}$, which means that $a_{m,n}$ is an indicator for how much interference U_m suffers at t_n. The following lemma shows an important property of $a_{m,n}$.

Lemma 2: Assume that $D_n - D_{n-1} \leq -\frac{N}{\log a_{n,1}}$, *for* $2 \leq n \leq M$, *and assume that the hybrid NOMA solution with* $\mathcal{S}_m = \{1, \cdots, m-1\}$ *is adopted in each step of Algorithm 1. Then, the following equality holds*

$$a_{m,i} = a_{m,l}, \tag{6}$$

for $i \neq l$, $1 \leq i \leq m-1$, *and* $1 \leq l \leq m-1$.

Proof: See Appendix B. □

Remark 7: Lemma 2 indicates that the hybrid NOMA power allocation solution is similar to the water-filling power allocation strategy [28]. In Fig. 2, a four-user example is used to illustrate this interesting property. U_2's transmit powers in t_1 and t_2 are allocated carefully; such that the interference levels (or water levels) seen by U_3 at t_1 and t_2 are the same, i.e., $a_{3,1} = a_{3,2}$. Furthermore, compared to t_2, U_2 experiences less interference in t_1, and hence uses more power in t_2. This is also analogous to the principle of water-filling, i.e., more water (i.e., transmit power) is poured to a hole which is deeper (i.e., t_2). Similarly, U_3 chooses its transmit powers to ensure that the interference experienced by U_4 in t_1, t_2, and t_3 reaches the same levels, i.e., $a_{4,1} = a_{4,2} = a_{4,3}$.

Intuitively, if U_4 suffers the same amount of interference in each of the first three time slots, it should use the same transmit power during these three time slots. Indeed, by combining Lemmas 1 and 2, this intuition regarding the users' transmit powers can be straightforwardly proved as shown in the following corollary.

[4]We note that the definition of $\mathcal{S}_m$ does not exclude pure NOMA offloading, i.e., U_m can still choose $P_{m,m} = 0$ and hence does not transmit during time slot t_m. We also note that U_m's data rates at t_m and t_i, $i < m$, are expressed differently, e.g., $R_{m,m} = \log\left(1 + P_{m,m}|h_m|^2\right)$ but $R_{m,i}$, $i < m$, is expressed differently as in (2), which is the reason why t_m is not included in $\mathcal{S}_m$.

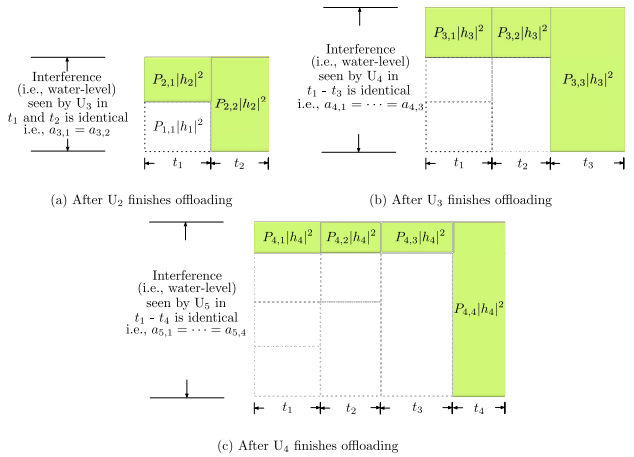


Fig. 2. Illustration of the water-filling-like power allocation strategy indicated by Lemma 2.

Corollary 1: Assuming that $D_n - D_{n-1} \leq -\frac{N}{\log a_{n,1}}$, *for* $2 \leq n \leq M$, *and assuming that the hybrid NOMA solution with* $\mathcal{S}_m = \{1, \cdots, m-1\}$ *is adopted in each step of Algorithm 1,* U_m*'s transmit powers are identical for all* t_i, $1 \leq i \leq m-1$, *i.e.,* $P_{m,i} = P_{m,l}$, *for* $i \neq l$, $1 \leq i \leq m-1$, *and* $1 \leq l \leq m-1$.

Remark 8: Lemma 2 can also be used to simplify the feasibility condition of hybrid NOMA shown in Lemma 1, i.e., $\min\{a_{m,1}, \cdots, a_{m,m-1}\} \geq e^{-\frac{N-\sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k \log a_{m,k}}{t_m + \sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k}}$ can be simplified to $D_m - D_{m-1} \leq -\frac{N}{\log a_{m,1}}$, where $t_m^* = D_m - D_{m-1}$ is used. Furthermore, Lemma 2 can be used to simplify the optimal solution of problem P3 by establishing the following two facts. The first compares the OMA and hybrid NOMA solutions, as stated in the following lemma.

Lemma 3: Assume that $D_n - D_{n-1} \leq -\frac{N}{\log a_{n,1}}$, *for* $2 \leq n \leq M$, *and* U_n, $2 \leq n \leq m-1$, *chooses the hybrid NOMA solution with* $\mathcal{S}_n = \{1, \cdots, n-1\}$. U_m*'s energy consumption in hybrid NOMA with* $\mathcal{S}_m = \{1, \cdots, m-1\}$ *is no more than that of the pure OMA solution.*

Proof: See Appendix C. □

The second fact is that among all possible hybrid NOMA strategies, U_m can reduce its energy consumption by choosing to transmit continuously, i.e., by adopting the hybrid NOMA solution with $\mathcal{S}_m = \{1, \cdots, m-1\}$, as shown in the following lemma.

Lemma 4: Assume that $D_n - D_{n-1} \leq -\frac{N}{\log a_{n,1}}$, *for* $2 \leq n \leq M$, *and assume that* U_n, $2 \leq n \leq m-1$, *chooses the hybrid NOMA solution with* $\mathcal{S}_n = \{1, \cdots, n-1\}$. *Thus, among all of* U_m*'s hybrid NOMA power allocation solutions shown in* (5), *the one with* $\mathcal{S}_m = \{1, \cdots, m\}$ *yields the lowest energy consumption.*

Proof: See Appendix D. □

Lemmas 3 and 4 indicate that U_m prefers the hybrid NOMA solution with $\mathcal{S}_m = \{1, \cdots, m-1\}$ if U_n, $n < m$, also chooses the same solution. By using a simple proof by mathematical induction and also the fact that U_2 chooses the hybrid NOMA solution whenever it is feasible which was established in [24], a more concise conclusion for the optimal solution of problem P3 can be straightforwardly obtained, as shown in the following corollary.

Corollary 2: Assume that $D_m - D_{m-1} \leq -\frac{N}{\log a_{m,1}}$, $2 \leq m \leq M$. *For the optimization problem shown in* (P3), *the optimal power allocation strategy is given by*

$$P_{m,n}^* = \begin{cases} \dfrac{e^{\frac{N-D_{m-1}\log a_{m,1}}{D_m}} - 1}{|h_m|^2}, & \text{if} \quad n = m \\ \dfrac{a_{m,n} e^{\frac{N-D_{m-1}\log a_{m,1}}{D_m}} - 1}{a_{m,1}|h_m|^2}, & \text{if} \quad n < m. \end{cases} \tag{7}$$

The optimal choice of t_m *is given by* $t_m^* = D_m - D_{m-1}$.

Remark 9: Compared to Lemma 1, Corollary 2 is much more concise and insightful. For example, Corollary 2 shows that there is a single optimal solution for problem P3, and the feasiblity condition for this hybrid NOMA based optimal solution is simply $D_m - D_{m-1} \leq -\frac{N}{\log a_{m,1}}$. For the two-user case, it was shown in [24] that the condition to switch from hybrid NOMA to OMA is $D_2 \geq 2D_1$. We note that the condition shown in Corollary 2 is consistent with that established for the two-user case. In particular, for $M = 2$, $D_m - D_{m-1} \leq -\frac{N}{\log a_{m,1}}$ is equivalent to $D_2 \geq D_1 + \frac{N}{\log(1+P_{1,1}|h_1|^2)} = 2D_1$ because $D_1 \log(1 + P_{1,1}|h_1|^2) = N$.

Furthermore, by using Lemma 2, the equivalence between problem P3 and problem P2 can be established, as shown in the following lemma.

Lemma 5: Assuming that $D_m - D_{m-1} \leq -\frac{N}{\log a_{m,1}}$, $2 \leq m \leq M$, *the solution shown in* (7) *is the optimal solution of problem P2.*

Proof: See Appendix E. □

Finally, the Pareto optimality of the obtained solution is shown in the following.

Lemma 6: Assuming that $D_m - D_{m-1} \leq -\frac{N}{\log a_{m,1}}$, $2 \leq m \leq M$, *the solution shown in* (7) *is a Pareto optimal solution of the multi-objective optimization problem shown in* (P1).

Proof: See Appendix F. □

Remark 10: As shown in the proof for Lemma 5, the energy consumption vector realized by the proposed algorithm, denoted by $\mathbf{E}_M^* \triangleq \left[E_2^* \cdots E_M^*\right]$, is a minimal element of all feasible $\mathbf{E}_M$ [29]. In other words, there is no feasible $\mathbf{E}_M$ which dominates $\mathbf{E}_M^*$, i.e., $\mathbf{E}_M \prec \mathbf{E}_M^*$. The Pareto optimality of the obtained solution will be investigated in detail in the next section by using the scalarization method.

IV. Numerical Studies

In this section, the performance of the proposed NOMA-MEC offloading scheme is studied via computer simulations.

A. Optimality of the Proposed Successive Optimization Algorithm

In Fig. 3, the optimal solution of problem P2 provided in Corollary 2 is verified by focusing on the three-user case and using an exhaustive search as a benchmark scheme. In particular, for the exhaustive search, U_2's transmission parameters are fixed by using the outcome of the first step of Algorithm 1,

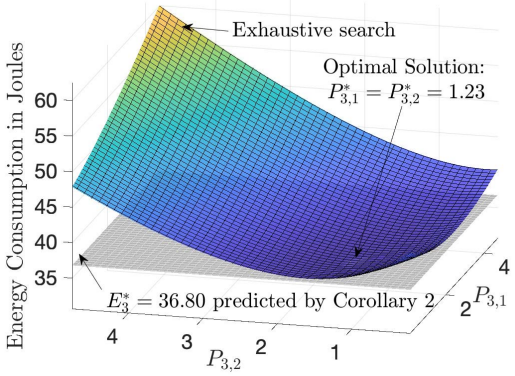


Fig. 3. Verification of the optimal solution shown in Corollary 2. $M = 3$ and $N = 10$, and the users' task deadlines are $D_1 = 8$, $D_2 = 12$, and $D_3 = 16$, respectively. For the exhaustive search, $P_{3,3}$ is obtained by using $P_{3,1}$ and $P_{3,2}$ to solve the following equation, $t_1 R_{3,1} + t_2 R_{3,2} + t_3 R_{3,3} = N$ and then checking whether the obtained $P_{3,3}$ satisfies the transmit power constraint in (P1d).

TABLE I

WEIGHTED ENERGY CONSUMPTION $\sum_{m=1}^{M} w_m E_m$ IN JOULES

	$N = 8$	$N = 10$	$N = 12$	$N = 14$	$N = 16$
$E^{\mathrm{w}}_{\mathrm{ES}}$ with $\mathbf{w}_1$	18.57	29.44	45.27	68.29	101.79
$E^{\mathrm{w}}_{\mathrm{SO}}$ with $\mathbf{w}_1$	18.57	29.44	45.27	68.29	101.79
$E^{\mathrm{w}}_{\mathrm{ES}} - E^{\mathrm{w}}_{\mathrm{SO}}$	1.07×10^{-4}	1.55×10^{-4}	1.03×10^{-5}	2.63×10^{-10}	1.64×10^{-6}
$E^{\mathrm{w}}_{\mathrm{ES}}$ with $\mathbf{w}_2$	19.66	31.59	49.19	75.11	113.21
$E^{\mathrm{w}}_{\mathrm{SO}}$ with $\mathbf{w}_2$	19.98	32.26	50.48	77.47	117.38
$E^{\mathrm{w}}_{\mathrm{ES}} - E^{\mathrm{w}}_{\mathrm{SO}}$	-0.32	-0.67	-1.29	-2.36	-4.17

and t_3 is also fixed as $t_3 = D_3 - D_2$. For illustration purposes, all users' channels are assumed to be normalized. As can be seen from Fig. 3, the minimum energy consumption predicted by Corollary 2 matches perfectly the exhaustive search result, which verifies the accuracy of Corollary 2.

The verification for the Pareto optimality of the obtained solution is based on the scalarization method which converts the multi-objective optimization problem into (P1) to the following single-objective optimization problem:

$$\text{(P4)}: \min_{\mathbf{x}} \quad \sum_{m=1}^{M} w_m E_m \quad \text{s.t. (P1b), (P1c), (P1d).}$$

Recall that different Pareto optimal solutions can be obtained by solving problem (P4) for different weights [29]. Therefore, in order to show that the solution obtained by the proposed algorithm is Pareto optimal, it is sufficient to show that the proposed solution can realize the same performance as the scalarization method with a particular set of weights. For Table I, two sets of the weights are used, namely $\mathbf{w}_1 = \left[\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3}\right]$ and $\mathbf{w}_2 = \frac{1}{6}\left[1\ 2\ 3\right]$, where $E^{\mathrm{w}}_{\mathrm{SO}}$ and $E^{\mathrm{w}}_{\mathrm{ES}}$ denote the weighted energy consumption achieved by the proposed algorithm and the scalarization method, respectively. Because problem (P4) is non-convex, an exhaustive search is used to solve this problem, where a step size of 0.1 is used. The

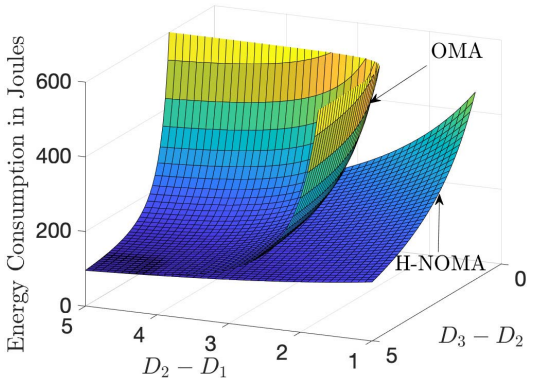


Fig. 4. Impact of the users' task deadlines on the energy consumed by MEC offloading with $M = 3$. $D_1 = 5$, and $N = 10$. H-NOMA refers to the proposed hybrid NOMA scheme.

simulation parameters used to generate Table I are the same as the ones used for Fig. 3. As can be seen from Table I, the proposed algorithm results in the same performance as the scalarization method for $\mathbf{w}_1$, which shows that the proposed solution is optimal for the case with $\mathbf{w}_1$. By using the fact that the solution to each scalarization yields a Pareto optimal solution, Table I confirms the Pareto-optimality of the solution obtained by the proposed algorithm [29]. For the case with $\mathbf{w}_2$, the solution obtained by the proposed algorithm results in a higher energy consumption than the scalarization method, which means that the proposed solution is suboptimal in this case. This is due to the fact that the solution developed in this paper is not a global optimal solution, but a Pareto-optimal solution only. The fact that the proposed algorithm can realize a Pareto-optimal solution is surprising, but might be due to the following reason. The proposed algorithm is similar to a greedy algorithm, where each user's transmission strategy is obtained separately by solving problem P2. As a result, the outcome of the proposed algorithm might be one of the stationary points of the considered problem, where no user has an incentive to deviate from its chosen strategy. The fact that $E^{\mathrm{w}}_{\mathrm{ES}} - E^{\mathrm{w}}_{\mathrm{SO}}$ is a small positive number instead of zero is explained in the following. For an optimization problem with continuous variables, the exhaustive search based scheme can only approximate the optimal solution. In other words, increasing the resolution of an exhaustive search reduces the gap between the outcome of the search and the optimal performance, but this gap cannot go to zero.

B. Performance of the Proposed Successive Optimization Algorithm

In this section, the performance of the proposed algorithm is studied by using OMA-MEC as a benchmark scheme, where the impact of different choices of the system parameters on MEC offloading is investigated. In Fig. 4, the overall energy consumption is shown as a function of the users' task deadlines, where $D_2 - D_1$ and $D_3 - D_2$ are used as the $x - y$ coordinates because they are the optimal choices for t_2 and t_3, respectively. As can be observed from Fig. 4, the use of hybrid NOMA-MEC can result in a significant

TABLE II

THE VALUES OF $-\frac{N}{\log a_{m,1}}$ FOR THE CURVES SHOWN IN FIG. 7

$D_m - D_{m-1}$	1	2	3	4	5	6	7	8	9	10
$m = 2$	5.50	6.00	6.50	7.00	7.50	8.00	8.50	9.00	9.50	10.00
$m = 3$	4.00	4.67	5.33	6.00	6.67	7.33	8.00	8.67	9.33	10.00
$m = 4$	3.25	4.00	4.75	5.50	6.25	7.00	7.75	8.50	9.25	10.00
$m = 3$	2.80	3.60	4.40	5.20	6.00	6.80	7.60	8.40	9.20	10.00

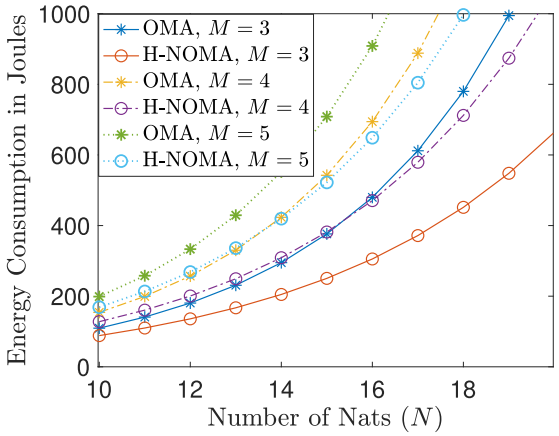


Fig. 5. Impact of the size of the tasks on the energy consumed by MEC offloading. $D_1 = 8$, and $D_m - D_{m-1} = \frac{D_1}{2}$.

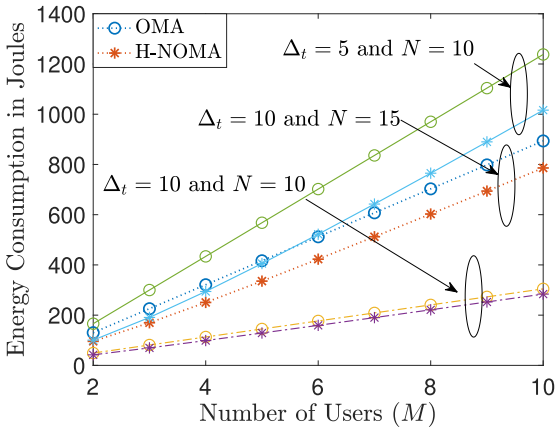


Fig. 6. Impact of the number of users on the energy consumed by MEC offloading. $D_1 = \Delta_t$, and $D_m - D_{m-1} = \frac{\Delta_t}{2}$.

reduction in energy consumption, compared to OMA-MEC. Furthermore, an interesting observation from Fig. 4 is that hybrid NOMA-MEC and OMA-MEC achieve the same performance for large $D_m - D_{m-1}$, which confirms Corollary 2, as for large $D_m - D_{m-1}$ the NOMA feasibility condition cannot be satisfied and hence hybrid NOMA-MEC degrades to OMA-MEC.

In Fig. 5, the impact of N on the energy consumption of MEC offloading is studied. As can be observed from the figure, the energy consumed by MEC offloading increases with N, which is due to the fact that a larger N means that more bits need to be offloaded and hence more energy needs to be consumed. For the considered range of N, the use of hybrid NOMA-MEC offers a reduction in energy consumption compared to OMA-MEC. Furthermore, Fig. 5 shows that the performance gain of NOMA-MEC over OMA-MEC is significantly larger for larger N, i.e., when the users' tasks are large, it is more beneficial to use hybrid NOMA-MEC.

In Fig. 6, the performance of MEC offloading is shown as a function of the number of users. In particular, the figure shows that increasing the number of users increases the overall energy consumption for both hybrid NOMA-MEC and OMA-MEC. Fig. 6 also shows that the two schemes require a higher energy consumption for MEC offloading, if a smaller $D_m - D_{m-1}$ is used, which can be explained by using OMA-MEC as an example. Recall that for OMA-MEC, U_m solely relies on t_m for offloading, and the optimal choice of t_m is $D_m - D_{m-1}$, which means that a smaller $D_m - D_{m-1}$ leads to less time available for offloading and hence more offloading energy is required. Fig. 6 shows that the performance gain of NOMA-MEC over OMA-MEC is larger if $D_m - D_{m-1}$ is smaller. This is expected since for OMA-MEC, U_m relies on $D_m - D_{m-1}$ (or t_m) for offloading, and a reduction of $D_m - D_{m-1}$ can cause a significant increase of energy consumption. For hybrid NOMA-MEC, U_m uses not only t_m but also t_n, $n < m$, for offloading, which makes hybrid NOMA-MEC less sensitive to the choice of $D_m - D_{m-1}$.

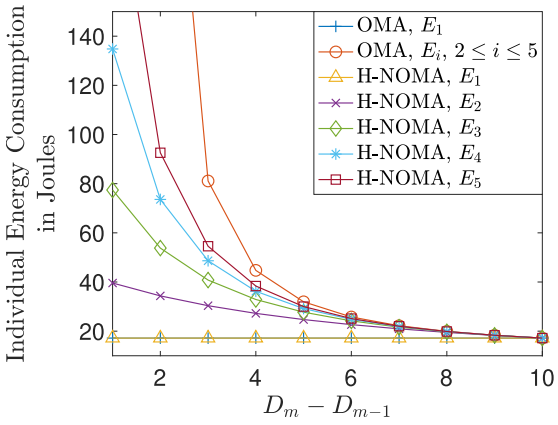


Fig. 7. Illustration of the condition under which hybrid NOMA-MEC and OMA-MEC achieve the same performance. $M = 5$, $N = 10$, and $D_1 = 5$. The differences between D_{m-1} and D_m are identical for $2 \le m \le M$.

Finally, the conditions under which hybrid NOMA-MEC and OMA-MEC yield the same performance are studied in Fig. 7, where the individual energy consumptions, E_m, are shown as functions of $D_m - D_{m-1}$. Recall that there are two possible optimal solutions for problem P2, one based on hybrid NOMA and one based on pure OMA. Remark 8 and Corollary 2 show that the feasibility of hybrid NOMA depends on the value of $D_m - D_{m-1}$, i.e., $D_n - D_{n-1} \le -\frac{N}{\log a_{n,1}}$. Therefore, a large value of $D_m - D_{m-1}$ can cause the situation, where the hybrid NOMA solution is infeasible and the OMA

solution is used. Fig. 7 and Table II confirm this conclusion, where increasing $D_m - D_{m-1}$ eventually leads to the situation, where hybrid NOMA-MEC is degraded to OMA-MEC. U_1's energy consumptions in the OMA and NOMA modes are the same, and E_1 is not a function of $D_m - D_{m-1}$ since $P_{1,1}^*$ and t_1^* are not functions of the other users' parameters. Among the other four users, U_i's energy consumption is less than U_j's, for $i < j$, since U_j suffers from more interference as illustrated in Fig. 2.

V. Conclusion

In this paper, a general hybrid NOMA-MEC offloading strategy has been proposed, and a multi-objective optimization problem has been formulated for minimization of the users' energy consumption for MEC offloading, where a low-complexity Pareto-optimal resource allocation solution has been obtained. Furthermore, by analyzing the properties of the obtained resource allocation solution, important insights regarding NOMA-MEC offloading have been obtained.

In this work, an exhaustive search has been used in order to find Pareto optimal solutions based on the scalarization method for the three-user case. When there are more than three users, exhaustive search can result in prohibitive computational complexity. Therefore, an important direction for future research is to develop a general and low complexity alternative to exhaustive search by applying advanced optimization tools, such as monotonic optimization. Furthermore, in this paper, the SIC decoding order is solely based on the users' quality of service (QoS) requirements, i.e., the users' task deadlines. The recent literature has demonstrated that the performance of uplink NOMA can be further improved by using the hybrid SIC decoding order that is based on both the users' QoS requirements and their channel conditions [30], [31]. Therefore, another important direction for future research is to investigate how to exploit hybrid SIC for further improving the performance of NOMA-MEC offloading.

Appendix A
Proof for Lemma 1

The lemma can be proved based on the following two steps. The first step is to fix t_m and find a closed-form expression for the optimal power allocation solution. The second step is to prove that $t_m^* = D_m - D_{m-1}$ for all possible power allocations.

A. Finding the Power Allocations

The Lagrangian of problem P3 can be expressed as follows:

$$L = P_{m,m}t_m + \sum_{n=1}^{m-1} P_{m,n}t_n - \sum_{n=1}^{m} \lambda_n P_{m,n} + \lambda_0 \left(N - t_m R_{m,m} - \sum_{n=1}^{m-1} t_n R_{m,n} \right), \quad (8)$$

where the λ_i's, $0 \le i \le m$, denote the Lagrange multipliers.

For a fixed t_m, it is straightforward to show that problem P3 is convex, and hence the optimal power allocation solutions can be obtained from solving the following KKT conditions:

$$\begin{cases} t_n - \lambda_0 t_n \dfrac{a_{m,n}|h_m|^2}{1 + a_{m,n}|h_m|^2 P_{m,n}} - \lambda_n = 0, \quad 1 \le n \le m-1 \\ t_m - \lambda_0 t_m \dfrac{|h_m|^2}{1 + |h_m|^2 P_{m,m}} - \lambda_m = 0 \\ \lambda_0 \left(t_m R_{m,m} + \sum_{n=1}^{m-1} t_n R_{m,n} - N \right) = 0 \\ \lambda_n P_{m,n} = 0, \quad 1 \le n \le m, \quad \text{(P2b), (P2c), \&(P3b),} \end{cases}$$

where the first two equations are obtained by using the Lagrangian shown in (8). The three considered types of transmission modes, namely pure OMA, pure NOMA, and hybrid NOMA, can be obtained as follows.

1) Pure OMA: Assume that $\lambda_n \neq 0$, $1 \le n \le m-1$, which means that $P_{m,n} = 0$, $1 \le n \le m-1$, and $\lambda_0 \neq 0$. Hence, $t_m R_{m,m} = N$, which yields the OMA solution shown in the lemma.

2) Hybrid NOMA: Recall that $\mathcal{S}_m$ includes the time slots during which U_m transmits. Therefore, for a hybrid NOMA solution, given $\mathcal{S}_m$, $\lambda_n = 0$ for $n \in \mathcal{S}_m$. Because $\lambda_n = 0$ for $n \in \mathcal{S}_m$, $\lambda_0 \neq 0$, which means that U_m definitely transmits at t_m. Because $\lambda_0 \neq 0$, $t_m R_{m,m} + \sum_{n=1, n\in\mathcal{S}_m}^{m-1} t_n R_{m,n} - N = 0$. Therefore, the KKT conditions can be simplified as follows:

$$\begin{cases} 1 - \lambda_0 \dfrac{a_{m,n}|h_m|^2}{1 + a_{m,n}|h_m|^2 P_{m,n}} = 0, \quad n \in \mathcal{S}_m \\ 1 - \lambda_0 \dfrac{|h_m|^2}{1 + |h_m|^2 P_{m,m}} = 0 \\ t_m \log\left(1 + P_{m,m}|h_m|^2\right) + \sum_{n=1, n\in\mathcal{S}_m}^{m-1} t_n \log\left(1 + a_{m,n}P_{m,n}|h_m|^2\right) = N, \end{cases} \quad (9)$$

which can be further simplified as follows:

$$\begin{cases} \dfrac{a_{m,n}}{1 + a_{m,n}|h_m|^2 P_{m,n}} = \dfrac{1}{1 + |h_m|^2 P_{m,m}}, \quad n \in \mathcal{S}_m \\ t_m \log\left(1 + P_{m,m}|h_m|^2\right) + \sum_{n=1, n\in\mathcal{S}_m}^{m-1} t_n \log\left(1 + a_{m,n}P_{m,n}|h_m|^2\right) = N. \end{cases} \quad (10)$$

Defining $y_n = \log\left(1 + a_{m,n}|h_m|^2 P_{m,n}\right)$ for $n \in \mathcal{S}_m$ and $y_m = \log\left(1 + |h_m|^2 P_{m,m}\right)$, the KKT conditions can be simplified as follows:

$$\begin{cases} \log a_{m,n} + y_m = y_n, \quad n \in \mathcal{S}_m \\ t_m y_m + \sum_{n=1, n\in\mathcal{S}_m}^{m-1} t_n y_n = N, \end{cases}$$

which yields the following solutions:

$$y_m = \frac{N - \sum_{k=1, k\in\mathcal{S}_m}^{m-1} t_k \log a_{m,k}}{t_m + \sum_{k=1, k\in\mathcal{S}_m}^{m-1} t_k},$$

and

$$y_n = \log a_{m,n} + \frac{N - \sum_{k=1, k\in\mathcal{S}_m}^{m-1} t_k \log a_{m,k}}{t_m + \sum_{k=1, k\in\mathcal{S}_m}^{m-1} t_k},$$

for $n \in \mathcal{S}_m$. By using y_m and y_n, $P_{m,m}^*$ and $P_{m,n}^*$ can be obtained as shown in the lemma, where the condition shown in the lemma, $a_{m,n} \ge e^{-\frac{N - \sum_{k=1, k\in\mathcal{S}_m}^{m-1} t_k \log a_{m,k}}{t_m + \sum_{k=1}^{m-1} t_k}}$, is required to

ensure that $P^*_{m,n} \geq 0$, $1 \leq n \leq m-1$. We note that $P^*_{m,n} \geq 0$ is a sufficient condition to ensure $P^*_{m,m} \geq 0$ because $a_{m,1} \leq 1$.

3) Pure NOMA: The pure NOMA solution, i.e., $P_{m,m} = 0$, can be discarded by treating it as a special case of hybrid NOMA and showing that it results in higher energy consumption than hybrid NOMA. Without loss of generality, we focus on the hybrid NOMA solution with $\mathcal{S}_m = \{1, \cdots, m-1\}$. The corresponding pure NOMA power allocation strategy can be obtained by letting $t_m \to 0$.

By using the hybrid NOMA solution shown in the lemma, the total energy consumption is given by

$$E_H = t_m \frac{e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - 1}{|h_m|^2} + \sum_{n=1}^{m-1} t_n \frac{a_{m,n} e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - 1}{a_{m,n}|h_m|^2}, \quad (11)$$

where the energy consumption of pure NOMA is also a special case of E_H by letting $t_m \to 0$.

By treating E_H as a function of t_m, i.e., $E_H(t_m)$, the superiority of hybrid NOMA over pure NOMA can be proved by showing $E_H(t_m) \leq E_H(0)$, as done in the following. The first oder derivative of E_H with respect to t_m is given by

$$E'_H = \frac{e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - 1}{|h_m|^2} - \frac{t_m}{|h_m|^2} e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} \times \frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{\left(t_m+\sum_{k=1}^{m-1} t_k\right)^2} - \sum_{n=1}^{m-1} \frac{a_{m,n} t_n}{a_{m,n}|h_m|^2} \times e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} \frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{\left(t_m+\sum_{k=1}^{m-1} t_k\right)^2}. \quad (12)$$

By defining $\theta = N - \sum_{k=1}^{m-1} t_k \log a_{m,k}$ and $\tau_m = \sum_{k=1}^{m} t_k$, the expression for the energy consumption can be simplified as follows:

$$\begin{aligned} |h_m|^2 E'_H &= e^{\frac{\theta}{t_m+\tau_{m-1}}} - 1 - t_m e^{\frac{\theta}{t_m+\tau_{m-1}}} \frac{\theta}{(t_m+\tau_{m-1})^2} \\ &\quad - \sum_{n=1}^{m-1} t_n e^{\frac{\theta}{t_m+\tau_{m-1}}} \frac{\theta}{(t_m+\tau_{m-1})^2} \\ &= e^{\frac{\theta}{t_m+\tau_{m-1}}} - 1 - e^{\frac{\theta}{t_m+\tau_{m-1}}} \frac{\theta}{t_m+\tau_{m-1}} \\ &= g\left(\frac{\theta}{t_m+\tau_{m-1}}\right), \end{aligned}$$

where $g(x) \triangleq e^x - 1 - xe^x$. In [24], it has been proved that $g(x)$ is a monotonically non-increasing function of x for $x \geq 0$, which means that $|h_m|^2 E'_H \leq g(0) = 0$. Therefore, the energy consumed by hybrid NOMA is a monotonically non-increasing function of t_m, i.e., $E_H(t_m) \leq E_H(0)$. Therefore, the energy consumption required by pure NOMA is no less than that of hybrid NOMA, and hence, the pure NOMA solution can be discarded.

B. Finding the Optimal Value for t_m

By using the closed-form expressions for the power allocation, problem P3 can be recasted as follows:

$$\min_{t_m} \quad P^*_{m,m} t_m + \sum_{n=1}^{m-1} P^*_{m,n} t_n \quad \text{(P5a)}$$

$$\text{s.t. } 0 \leq t_m \leq D_m - \sum_{n=1}^{m-1} t_n. \quad \text{(P5b)}$$

In problem P5, either the power allocation strategy for pure OMA or that for hybrid NOMA can be used. For the case of hybrid NOMA, in the previous section, it was proved that the energy consumption for hybrid NOMA, E_H, is a monotonically non-increasing function of t_m, which means that the optimal choice of t_m is $t^*_m = D_m - \sum_{n=1}^{m-1} t_n$.

For the OMA case, the total energy consumption is given by

$$E_O = t_m \frac{e^{\frac{N}{t_m}} - 1}{|h_m|^2}. \quad (13)$$

The first order derivative of E_O with respect to t_m is given by

$$E'_O = \frac{e^{\frac{N}{t_m}} - 1 - \frac{N}{t_m} e^{\frac{N}{t_m}}}{|h_m|^2} = \frac{g\left(\frac{N}{t_m}\right)}{|h_m|^2}. \quad (14)$$

Exploiting again the fact that $g(x)$ is a monotonically non-increasing function of x, the energy consumed by pure OMA is a monotonically non-increasing function of t_m, and hence $t^*_m = D_m - \sum_{n=1}^{m-1} t_n$, which is the same as for hybrid NOMA.

Finally, because of the successive nature of Algorithm 1, i.e., $t^*_n = D_n - D_{n-1}$, for $2 \leq n < m$, and hence $t^*_m = D_m - D_{m-1}$. The proof of the lemma is complete.

APPENDIX B
PROOF FOR LEMMA 2

Recall that $a_{m,i}$ is defined as follows:

$$a_{m,i} = \frac{1}{\sum_{j=i}^{m-1} P_{j,i}|h_j|^2 + 1}. \quad (15)$$

By using the hybrid NOMA power allocation solutions, $a_{m,i}$ can be first expressed as follows:

$$\begin{aligned} \frac{1}{a_{m,i}} &= \sum_{j=i}^{m-1} P_{j,i}|h_j|^2 + 1 = P_{i,i}|h_i|^2 + \sum_{j=i+1}^{m-1} P_{j,i}|h_j|^2 + 1 \\ &= e^{\frac{N-\sum_{k=1}^{i-1} t_k \log a_{i,k}}{t_i+\sum_{k=1}^{i-1} t_k}} + \sum_{j=i+1}^{m-1} \frac{a_{j,i} e^{\frac{N-\sum_{k=1}^{j-1} t_k \log a_{j,k}}{t_j+\sum_{k=1}^{j-1} t_k}} - 1}{a_{j,i}}, \end{aligned} \quad (16)$$

for $1 \leq i \leq m-2$, and

$$\frac{1}{a_{m,m-1}} = e^{\frac{N-\sum_{k=1}^{m-2} t_k \log a_{m-1,k}}{t_{m-1}+\sum_{k=1}^{m-2} t_k}}. \quad (17)$$

The lemma can be proved by mathematical induction. The smallest value for m in the lemma is $m = 3$. Therefore, the proof is divided into two parts. The first part is about the base

case, where $\frac{1}{a_{3,2}} - 1 = \frac{1}{a_{3,1}} - 1$ is proved. The second part is for the inductive step, where, by assuming that the lemma holds for $m-1$, the lemma for the case of m is proved.

A. The Base Case $m = 3$

In this part, we aim to prove the equivalence of $a_{3,1}$ and $a_{3,2}$. For the case of $m = 3$, (17) can be expressed as follows:

$$\frac{1}{a_{3,2}} = e^{\frac{N - t_1 \log a_{2,1}}{t_2 + t_1}}. \tag{18}$$

On the other hand, $a_{3,1}$ can be written as follows:

$$\begin{aligned}\frac{1}{a_{3,1}} &= e^{\frac{N}{t_1}} + \frac{a_{2,1} e^{\frac{N - t_1 \log a_{2,1}}{t_2 + t_1}} - 1}{a_{2,1}} \\ &= e^{\frac{N}{t_1}} + e^{\frac{N - t_1 \log a_{2,1}}{t_2 + t_1}} - \frac{1}{a_{2,1}}.\end{aligned} \tag{19}$$

Both the expressions for $a_{3,1}$ and $a_{3,2}$ are related to $a_{2,1}$. Recall that $a_{2,1} = \frac{1}{P_{1,1}|h_1|^2 + 1}$, where $P_{1,1}$ is the transmit power used by U_1 at t_1. Because of the constraint $t_1 \log(1 + P_{1,1}|h_1|^2) = N$, the following equality holds:

$$e^{\frac{N}{t_1}} = \frac{1}{a_{2,1}}, \tag{20}$$

which means that (19) can be simplified as follows:

$$\frac{1}{a_{3,1}} = e^{\frac{N - t_1 \log a_{2,1}}{t_2 + t_1}}. \tag{21}$$

Combining (18) and (21), one can conclude that the lemma holds for the case of $m = 3$, i.e.,

$$a_{3,1} = a_{3,2}. \tag{22}$$

B. Inductive Step

Assume now that the lemma holds for the cases of i, $3 \leq i \leq m-1$, i.e.,

$$a_{i,1} = \cdots = a_{i,i-1}, \quad 3 \leq i \leq m-1. \tag{23}$$

The aim of this part is to show that the lemma also holds for the case of m, i.e.,

$$a_{m,1} = \cdots = a_{m,m-1}. \tag{24}$$

First rewrite $a_{m,i}$, $1 \leq i \leq m-2$, shown in (16) as follows:

$$\begin{aligned}\frac{1}{a_{m,i}} = {} & e^{\frac{N - \sum_{k=1}^{m-2} t_k \log a_{m-1,k}}{t_{m-1} + \sum_{k=1}^{m-2} t_k}} \\ & + \sum_{j=i}^{m-2} e^{\frac{N - \sum_{k=1}^{j-1} t_k \log a_{j,k}}{t_j + \sum_{k=1}^{j-1} t_k}} - \sum_{j=i+1}^{m-1} \frac{1}{a_{j,i}}.\end{aligned} \tag{25}$$

Comparing (25) to (17), one can observe that proving (24) is equivalent to proving the following:

$$\breve{\Delta} \triangleq \sum_{j=i}^{m-2} e^{\frac{N - \sum_{k=1}^{j-1} t_k \log a_{j,k}}{t_j + \sum_{k=1}^{j-1} t_k}} - \sum_{j=i+1}^{m-1} \frac{1}{a_{j,i}} = 0, \tag{26}$$

for any $1 \leq i \leq m-2$. $\breve{\Delta}$ can be further rewritten as follows:

$$\begin{aligned}\breve{\Delta} &= \sum_{j=i}^{m-2} \left(e^{\frac{N - \sum_{k=1}^{j-1} t_k \log a_{j,k}}{t_j + \sum_{k=1}^{j-1} t_k}} - \frac{1}{a_{j+1,i}} \right) \\ &= \sum_{j=i}^{m-2} \left(e^{\frac{N - \left(\sum_{k=1}^{j-1} t_k\right) \log a_{j,1}}{t_j + \sum_{k=1}^{j-1} t_k}} - \frac{1}{a_{j+1,1}} \right),\end{aligned} \tag{27}$$

where the last step follows from the assumption in (23).

Define the difference between the two terms in the bracket of (27) as Δ, which can be expressed as follows:

$$\begin{aligned}\Delta &= e^{\frac{N - \left(\sum_{k=1}^{j-1} t_k\right) \log a_{j,1}}{t_j + \sum_{k=1}^{j-1} t_k}} - \frac{1}{a_{j+1,1}} \\ &= e^{\frac{N - \left(\sum_{k=1}^{j-1} t_k\right) \log a_{j,1}}{t_j + \sum_{k=1}^{j-1} t_k}} - e^{-\log a_{j+1,1}},\end{aligned} \tag{28}$$

which means that $\Delta = 0$ if the difference of the exponents of the two exponential functions in (28), denoted by $\bar{\Delta}$, is zero. We note that $\bar{\Delta}$ can be expressed as follows:

$$\begin{aligned}\bar{\Delta} &= \frac{N - \left(\sum_{k=1}^{j-1} t_k\right) \log a_{j,1}}{t_j + \sum_{k=1}^{j-1} t_k} + \log a_{j+1,1} \\ &= \frac{N - \left(\sum_{k=1}^{j-1} t_k\right) \log a_{j,1} + \left(t_j + \sum_{k=1}^{j-1} t_k\right) \log a_{j+1,1}}{t_j + \sum_{k=1}^{j-1} t_k}.\end{aligned} \tag{29}$$

Recall that $a_{m,n}$ can also be related to the users' data rates. Therefore, $\bar{\Delta}$ can be further expressed as follows:

$$\begin{aligned}\bar{\Delta} &= \frac{N + t_j \log a_{j+1,1} + \sum_{k=1}^{j-1} t_k \left(\log a_{j+1,1} - \log a_{j,1}\right)}{t_j + \sum_{k=1}^{j-1} t_k} \\ &= \frac{N + t_j \log a_{j+1,j} + \sum_{k=1}^{j-1} t_k \left(\log a_{j+1,k} - \log a_{j,k}\right)}{t_j + \sum_{k=1}^{j-1} t_k},\end{aligned}$$

where the last step follows from the assumption in (23).

On the one hand, $\log a_{j+1,j}$ can be related to U_j's offloading data rate at t_j as follows:

$$\log a_{j+1,j} = -\log\left(P_{j,j}|h_j|^2 + 1\right) = -R_{j,j}. \tag{30}$$

On the other hand, $\log a_{j+1,k} - \log a_{j,k}$ can be related to U_j's offloading data rate at t_k as follows:

$$\begin{aligned}\log a_{j+1,k} - \log a_{j,k} &= -\log \frac{\sum_{p=k}^{j} P_{p,k}|h_p|^2 + 1}{\sum_{p=k}^{j-1} P_{p,k}|h_p|^2 + 1} \\ &= -\log\left(1 + \frac{P_{j,k}|h_j|^2}{\sum_{p=k}^{j-1} P_{p,k}|h_p|^2 + 1}\right) \\ &= -R_{j,k}.\end{aligned} \tag{31}$$

So the difference $\bar{\Delta}$ can be finally expressed as follows:

$$\bar{\Delta} = \frac{N - R_{j,j} - \sum_{k=1}^{j-1} R_{j,k}}{t_i + \sum_{k=1}^{j-1} t_k} = 0, \tag{32}$$

where the last step follows from the fact that, with the used hybrid NOMA power allocation, U_j can successfully finish its offloading. Because $\bar{\Delta} = 0$, $\breve{\Delta} = 0$, and hence, $a_{m,1} = \cdots = a_{m,m-1}$, i.e., the lemma holds for the case of m.

By combining the conclusions from the base case and the inductive step, the lemma is proved.

APPENDIX C
PROOF FOR LEMMA 3

Recall that the OMA power allocation solution is $P_{m,m}^{O*} = \frac{e^{\frac{N}{t_m}}-1}{|h_m|^2}$, and hence the energy consumed by OMA offloading is given by $E_O = t_m \frac{e^{\frac{N}{t_m}}-1}{|h_m|^2}$. The difference between the energy consumed by hybrid NOMA and OMA, denoted by $\tilde{\Delta} \triangleq |h_m|^2(E_H - E_O)$, is given by

$$\tilde{\Delta} = t_m e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} + \sum_{n=1}^{m-1} t_n \frac{a_{m,n} e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - 1}{a_{m,n}} - t_m e^{\frac{N}{t_m}}. \quad (33)$$

The lemma is proved by showing that $\tilde{\Delta} \leq 0$ for $a_{m,n} e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - 1 \geq 0$.

The difference $\tilde{\Delta}$ can be further rewritten as follows:

$$\begin{aligned}\tilde{\Delta} &= t_m e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} + \sum_{n=1}^{m-1} t_n e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - \sum_{n=1}^{m-1} t_n \frac{1}{a_{m,n}} - t_m e^{\frac{N}{t_m}} \\ &= \sum_{n=1}^{m} t_n e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - \sum_{n=1}^{m-1} t_n \frac{1}{a_{m,n}} - t_m e^{\frac{N}{t_m}}. \quad (34)\end{aligned}$$

By applying Lemma 2, the difference $\tilde{\Delta}$ can be simplified as follows:

$$\tilde{\Delta} = e^{\frac{N-\left(\sum_{k=1}^{m-1} t_k\right) \log a_{m,1}}{t_m+\sum_{k=1}^{m-1} t_k}} \sum_{n=1}^{m} t_n - \frac{1}{a_{m,1}} \sum_{n=1}^{m-1} t_n - t_m e^{\frac{N}{t_m}}. \quad (35)$$

Recall that $\tau_m = \sum_{k=1}^{m} t_k$, which can be used to simplify $\tilde{\Delta}$ as follows:

$$\tilde{\Delta} = \tau_m e^{\frac{N-\tau_{m-1} \log a_{m,1}}{\tau_m}} - \tau_{m-1} \frac{1}{a_{m,1}} - t_m e^{\frac{N}{t_m}}. \quad (36)$$

We note that by using Lemma 2, the feasibility condition of hybrid NOMA, $\min\{a_{m,1}, \cdots, a_{m,m-1}\} \geq e^{-\frac{N-\sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1,k\in\mathcal{S}_m}^{m-1} t_k}}$, can be simplified as follows:

$$a_{m,1}^{-1} \leq e^{\frac{N}{t_m}}. \quad (37)$$

Because the feasibility condition is a lower bound on $a_{m,1}$, $\tilde{\Delta}$ in (36) can be expressed as an explicit function of $a_{m,1}$ as follows:

$$\tilde{\Delta} = \tau_m e^{\frac{N}{\tau_m}} \left(a_{m,1}^{-1}\right)^{\frac{\tau_{m-1}}{\tau_m}} - \tau_{m-1}\left(a_{m,1}^{-1}\right) - t_m e^{\frac{N}{t_m}} = f(a_{m,1}^{-1}), \quad (38)$$

where $f(x) \triangleq \tau_m e^{\frac{N}{\tau_m}} x^{\frac{\tau_{m-1}}{\tau_m}} - \tau_{m-1} x - t_m e^{\frac{N}{t_m}}$.

The lemma can be proved by showing $f(x) \leq 0$ for $x \leq e^{\frac{N}{t_m}}$. The first order derivative of $f(x)$ is given by

$$\begin{aligned}f'(x) &= \tau_m e^{\frac{N}{\tau_m}} \left(\frac{\tau_{m-1}}{\tau_m}\right) x^{\frac{\tau_{m-1}}{\tau_m}-1} - \tau_{m-1} \\ &= \tau_{m-1} e^{\frac{N}{\tau_m}} x^{-\frac{t_m}{\tau_m}} - \tau_{m-1}, \quad (39)\end{aligned}$$

which is a monotonically decreasing function of x. Therefore, for $x \leq e^{\frac{N}{t_m}}$, $f'(x)$ is lower bounded as follows:

$$f'(x) \geq \tau_{m-1} e^{\frac{N}{\tau_m}} \left(e^{\frac{N}{t_m}}\right)^{-\frac{t_m}{\tau_m}} - \tau_{m-1} \geq 0, \quad (40)$$

which means that $f(x)$ is a monotonically increasing function of x. Therefore, for $x \leq e^{\frac{N}{t_m}}$, $f(x)$ is upper bounded as follows:

$$\begin{aligned}f(x) &\leq f\left(e^{\frac{N}{t_m}}\right) = \tau_m e^{\frac{N}{\tau_m}} e^{\frac{N\tau_{m-1}}{t_m\tau_m}} - \tau_{m-1} e^{\frac{N}{t_m}} - t_m e^{\frac{N}{t_m}} \\ &= \tau_m \left(e^{\frac{N}{\tau_m}+\frac{N\tau_{m-1}}{t_m\tau_m}} - e^{\frac{N}{t_m}}\right). \quad (41)\end{aligned}$$

Define the difference of the exponents of the two exponential functions in (41) as $\breve{\Delta}$, which can be evaluated as follows:

$$\breve{\Delta} \triangleq \frac{N}{\tau_m} + \frac{N\tau_{m-1}}{t_m\tau_m} - \frac{N}{t_m} = \frac{N\tau_m - N\tau_m}{t_m\tau_m} = 0.$$

This means that

$$f(x) \leq f\left(e^{\frac{N}{t_m}}\right) = 0.$$

Therefore, $\tilde{\Delta} \leq 0$, i.e., the energy consumption of OMA offloading is no less than that of hybrid NOMA. The proof of the lemma is completed.

APPENDIX D
PROOF FOR LEMMA 4

The lemma can be proved by using the hybrid NOMA solution with $\tilde{\mathcal{S}}_m = \{2, \cdots, m-1\}$ as an example and showing that it does not consume less energy than the hybrid NOMA solution with $\mathcal{S}_m = \{1, \cdots, m-1\}$. Recall that the energy consumed by the hybrid NOMA solution with $\mathcal{S}_m = \{1, \cdots, m-1\}$ is given by

$$E_H = t_m \frac{e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - 1}{|h_m|^2} + \sum_{n=1}^{m-1} t_n \frac{a_{m,n} e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - 1}{a_{m,n}|h_m|^2}, \quad (42)$$

where the energy consumed by the solution with $\tilde{\mathcal{S}}_m = \{2, \cdots, m-1\}$, denoted by $\tilde{E}_H$, is a special case of E_H by letting $t_1 \to 0$. Therefore, in order to prove $\tilde{E}_H \geq E_H$, it is sufficient to show that E_H is a monotonically non-increasing function of t_1. The first order derivative of E_H with respect

to t_1 is given by

$$E'_H = \sum_{n=1}^{m} \frac{t_n}{|h_m|^2} e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} \times \left(\frac{-\log a_{m,1}}{t_m+\sum_{k=1}^{m-1} t_k} - \frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{\left(t_m+\sum_{k=1}^{m-1} t_k\right)^2} \right) + \frac{a_{m,1} e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - 1}{a_{m,1}|h_m|^2}. \quad (43)$$

Because E_H is achieved by the hybrid NOMA solution with $\mathcal{S}_m = \{1, \cdots, m-1\}$, Lemma 2 can be applied to simplify E'_H as follows:

$$\begin{aligned} E'_H &= \frac{1}{|h_m|^2} e^{\frac{N-\tau_{m-1}\log a_{m,1}}{\tau_m}} \left(-\frac{N+t_m \log a_{m,1}}{\tau_m} \right) + \frac{a_{m,1} e^{\frac{N-\tau_{m-1}\log a_{m,1}}{\tau_m}} - 1}{a_{m,1}|h_m|^2} \\ &= \frac{1}{|h_m|^2} e^{\frac{N-\tau_{m-1}\log a_{m,1}}{\tau_m}} \left(\frac{\tau_m - N - t_m \log a_{m,1}}{\tau_m} \right) - \frac{1}{|h_m|^2} e^{-\log a_{m,1}} = \frac{1}{|h_m|^2} p\left(-\log a_{m,1}\right), \end{aligned}$$

where $p(x) \triangleq e^{\frac{N+\tau_{m-1}x}{\tau_m}} \left(\frac{\tau_m - N + t_m x}{\tau_m} \right) - e^x$.

Recall that the feasibility condition for hybrid NOMA has been provided in (37) and it can be further rewritten as follows:

$$-\log a_{m,1} \le \frac{N}{t_m}. \quad (44)$$

Therefore, the fact that $E'_H \le 0$ can be proved by proving that $p(x) \le 0$ for $x \le \frac{N}{t_m}$, as shown in the following. Rewrite $p(x)$ as follows:

$$p(x) = e^{\frac{N+\tau_{m-1}x}{\tau_m}} e^{\log\left(\frac{\tau_m - N + t_m x}{\tau_m}\right)} - e^x. \quad (45)$$

Define $\tilde{p}(x) \triangleq \frac{N+\tau_{m-1}x}{\tau_m} + \log\left(\frac{\tau_m - N + t_m x}{\tau_m}\right) - x$. We note that $p(x) \le 0$ is equivalent to $\tilde{p}(x) \le 0$. The first order derivative of $\tilde{p}(x)$ is given by

$$\tilde{p}'(x) = \frac{\tau_{m-1}}{\tau_m} + \frac{t_m}{\tau_m} \frac{\tau_m}{\tau_m - N + t_m x} - 1. \quad (46)$$

We note that $\tilde{p}'(x)$ is a monotonically non-increasing function of x, which means

$$\tilde{p}'(x) \ge \tilde{p}'\left(\frac{N}{t_m}\right) = \frac{-t_m}{\tau_m} + \frac{t_m}{\tau_m} \frac{\tau_m}{\tau_m - N + N} = 0, \quad (47)$$

since $x \le \frac{N}{t_m}$. Therefore, $\tilde{p}(x)$ is a monotonically non-decreasing function of x, which means that the use of the upper bound on x, $x \le \frac{N}{t_m}$, yields the following upper bound on $p(x)$:

$$p(x) \le p\left(\frac{N}{t_m}\right) = e^{\frac{N+\tau_{m-1}\frac{N}{t_m}}{\tau_m}} e^{\log\left(\frac{\tau_m - N + N}{\tau_m}\right)} - e^{\frac{N}{t_m}} = 0. \quad (48)$$

Therefore, $E'_H \le 0$, and hence E_H is a monotonically non-increasing function of t_1, which means that the hybrid NOMA solution without using t_1 results in no less energy consumption than the solution with $\mathcal{S}_m = \{1, \cdots, m-1\}$. The proof is complete.

Appendix E
Proof for Lemma 5

The lemma can be proved by showing that the power allocation solutions of problem P3 do not violate the transmit power constraint of problem P2. Since P_m^{OMA} is the transmit power which is sufficient for U_m to complete OMA offloading by using t_m only, one can conclude that $P^*_{m,m} \le P_m^{\mathrm{OMA}}$, which means that the lemma can be proved by showing that $P^*_{m,n} \le P_m^{\mathrm{OMA}}$, $n < m$.

Recall that $P^*_{m,n}$, $n < m$, can be expressed as follows:

$$P^*_{m,n} = \frac{a_{m,n} e^{\frac{N-\sum_{k=1}^{m-1} t_k \log a_{m,k}}{t_m+\sum_{k=1}^{m-1} t_k}} - 1}{a_{m,n}|h_m|^2} = \frac{a_{m,1} e^{\frac{N-\tau_{m-1}\log a_{m,1}}{\tau_m}} - 1}{a_{m,1}|h_m|^2}, \quad (49)$$

where the last step follows from Lemma 2. Therefore, the difference between $P^*_{m,n}$ and the maximal power P_m^{OMA} is given by

$$\Delta_m \triangleq P^*_{m,n} - P_m^{\mathrm{OMA}} = \frac{a_{m,1} e^{\frac{N-\tau_{m-1}\log a_{m,1}}{\tau_m}} - 1}{a_{m,1}|h_m|^2} - P_m^{\mathrm{OMA}} = \frac{a_{m,1} e^{\frac{N-\tau_{m-1}\log a_{m,1}}{\tau_m}} - 1 - P_m^{\mathrm{OMA}} a_{m,1}|h_m|^2}{a_{m,1}|h_m|^2},$$

such that the lemma can be proved by showing that $\Delta_m \le 0$.

Recall that, in OMA, P_m^{OMA} is used to deliver N nats by using t_m only, i.e., $t_m \log(1 + P_m^{\mathrm{OMA}}|h_m|^2) = N$, which can be used to rewrite the difference as follows:

$$\Delta_m = \frac{a_{m,1} e^{\frac{N-\tau_{m-1}\log a_{m,1}}{\tau_m}} - 1 - a_{m,1}\left(e^{\frac{N}{t_m}} - 1\right)}{a_{m,1}|h_m|^2}. \quad (50)$$

Since $a_{m,1} \le 1$, $\Delta_m \le 0$ can be proved by showing the following inequality

$$\bar{\Delta}_m \triangleq \frac{N-\tau_{m-1}\log a_{m,1}}{\tau_m} - \frac{N}{t_m} \le 0. \quad (51)$$

$\bar{\Delta}_m$ can be further rewritten as follows:

$$\bar{\Delta}_m = \frac{t_m N - t_m \tau_{m-1} \log a_{m,1} - \tau_m N}{\tau_m t_m} = \frac{-t_m \tau_{m-1} \log a_{m,1} - \tau_{m-1} N}{\tau_m t_m}, \quad (52)$$

where it is important to point out that $-t_m \tau_{m-1} \log a_{m,1}$ is non-negative since $a_{m,1} \le 1$.

We recall that the feasibility condition for the hybrid NOMA solution is $t_m \le -\frac{N}{\log a_{m,1}}$, which means that the difference can be upper bounded as follows:

$$\bar{\Delta}_m \le \frac{\frac{N}{\log a_{m,1}} \tau_{m-1} \log a_{m,1} - \tau_{m-1} N}{\tau_m t_m} = 0, \quad (53)$$

which means that $P^*_{m,n} \le P_m^{\mathrm{OMA}}$, $n < m$. Therefore, the proof of the lemma is complete.

APPENDIX F
PROOF FOR LEMMA 6

Define $\mathcal{E}_M$ as a set collecting all feasible $\mathbf{E}_M$. The lemma is equivalent to the statement that there is no vector in $\mathcal{E}_M$ dominating $\mathbf{E}_M^*$. The lemma can be proved by using again mathematical induction.

For the base case, $M = 2$, a conclusion stronger than the lemma can be proved. In particular, we will show that $\mathbf{E}_2^*$ is the minimum element of $\mathcal{E}_2$, i.e., $\mathbf{E}_2^* < \mathbf{E}_2$ or equivalently $E_2^* < \mathbf{E}_2(1)$, for any $\mathbf{E}_2 \in \mathcal{E}_2$, where $\mathbf{E}_m(i)$ denotes the i-th element of the $(m-1)$ vector, $\mathbf{E}_m$. Recall that E_2^* is obtained by solving the following optimization problem

$$\min_{t_2, P_{2,n}} \quad E_2 \tag{P6a}$$

$$\text{s.t. } t_1^* R_{2,1} + t_2 R_{2,2} \geq N, \quad t_2 \leq D_2 - D_1, \quad t_2 \geq 0, \tag{P6b}$$

$$0 \leq P_{2,n} \leq P_2^{\text{OMA}}, \quad 1 \leq n \leq 2. \tag{P6c}$$

Because problem P6 is obtained from problem P1 by discarding the constraints related to U_i's parameters, $i > 2$, the optimal value of problem P6 is an achievable lower bound on the optimal value of problem P1. In other words, $E_2^* \leq \mathbf{E}_2(1)$, for any $\mathbf{E}_2 \in \mathcal{E}_2$. Therefore, the lemma holds for the base case.

For the inductive step, assume that the lemma holds for the case with $(m-1)$ users, i.e., $\mathbf{E}_{m-1}^*$ is Pareto optimal. In order to prove that $\mathbf{E}_m^*$ is also Pareto optimal, a proof by contradiction is used. In particular, assume that the lemma does not hold for the case with m users, i.e., there exists a vector in $\mathcal{E}_m$ which is denoted by $\bar{\mathbf{E}}_m$ and satisfies $\bar{\mathbf{E}}_m \prec \mathbf{E}_m^*$. Because of the assumption that $\mathbf{E}_{m-1}^*$ is Pareto optimal, $\bar{\mathbf{E}}_m \prec \mathbf{E}_m^*$ means that $\bar{\mathbf{E}}_m(m) < E_m^*$, and $\bar{\mathbf{E}}_m(i) = E_i^*$ for $1 \leq i \leq m-1$.

Since E_m^* is the optimal value of problem P2, $\bar{\mathbf{E}}_m(m) < E_m^*$ is possible only if the following event happens. In particular, one or multiple users choose their transmission parameters different from $P_{m,n}^*$ and t_n^*, for $1 \leq n \leq m-1$, which changes the feasibility set of problem P2. Among these users, denote by $\text{U}_{\bar{n}}$, $2 \leq \bar{n} \leq n$, the user whose signal is decoded first, which leads to the following two conclusions. The first is that $\bar{\mathbf{E}}_m(\bar{n}) = E_{\bar{n}}^*$, and the second is that the users, U_i, $1 \leq i \leq \bar{n}-1$, choose t_i^* and $P_{i,j}^*$, $1 \leq j \leq i$, for their transmission. The statement that $\text{U}_{\bar{n}}$ did not choose $t_{\bar{n}}^*$ and $P_{\bar{n},j}^*$, $1 \leq j \leq \bar{n}$, but still realizes $E_{\bar{n}}^*$ is equivalent to the statement that, for the following optimization problem, there are two different optimal solutions realizing the same optimal value:

$$\min_{\mathbf{p}_{\bar{n}}, t_{\bar{n}}} \quad E_{\bar{n}} \tag{P7a}$$

$$\text{s.t. } \sum_{j=1}^{\bar{n}} t_j^* R_{\bar{n},j} + t_{\bar{n}} R_{\bar{n},\bar{n}} \geq N, \tag{P7b}$$

$$t_{\bar{n}} \leq D_{\bar{n}} - \sum_{j=1}^{\bar{n}-1} t_j, \quad t_{\bar{n}} \geq 0, \tag{P7c}$$

$$0 \leq P_{\bar{n},j} \leq P_{\bar{n}}^{\text{OMA}}, \quad 1 \leq j \leq \bar{n}. \tag{P7d}$$

However, Corollary 2 shows that there is a single optimal solution for problem P7. Therefore, there cannot exist an $\bar{\mathbf{E}}_m$ in $\mathcal{E}_m$ which satisfies $\bar{\mathbf{E}}_m \prec \mathbf{E}_m^*$, which completes the inductive step.

By combining the conclusions from the base case and the inductive step, the lemma is proved.

REFERENCES

[1] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, Dec. 2017.

[2] Q.-V. Pham *et al.*, "A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art," *IEEE Access*, vol. 8, pp. 116974–117017, 2020.

[3] J. Santos, T. Wauters, B. Volckaert, and F. De Turck, "Towards low-latency service delivery in a continuum of virtual resources: State-of-the-art and research directions," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 4, pp. 2557–2589, Jul. 2021.

[4] T. K. Rodrigues, K. Suto, and N. Kato, "Edge cloud server deployment with transmission power control through machine learning for 6G Internet of Things," *IEEE Trans. Emerg. Topics Comput.*, vol. 9, no. 4, pp. 2099–2108, Oct. 2021.

[5] T. K. Rodrigues, J. Liu, and N. Kato, "Offloading decision for mobile multi-access edge computing in a multi-tiered 6G network," *IEEE Trans. Emerg. Topics Comput.*, early access, Jun. 17, 2021, doi: 10.1109/TETC.2021.3090061.

[6] H. Li, K. Ota, and M. Dong, "Learning IoT in edge: Deep learning for the Internet of Things with edge computing," *IEEE Netw.*, vol. 32, no. 1, pp. 96–101, Jan. 2018.

[7] W. Y. B. Lim *et al.*, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, 3rd Quart., 2020.

[8] Z. Ding, R. Schober, and H. V. Poor, "No-pain no-gain: DRL assisted optimization in energy-constrained CR-NOMA networks," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 5917–5932, Sep. 2021.

[9] M. Elbayoumi *et al.*, "NOMA-assisted machine-type communications in UDN: State-of-the-art and challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1276–1304, 2nd Quart., 2020.

[10] O. Maraqa, A. S. Rajasekaran, S. Al-Ahmadi, H. Yanikomeroglu, and S. M. Sait, "A survey of rate-optimal power domain NOMA with enabling technologies of future wireless networks," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2192–2235, 4th Quart., 2020.

[11] Z. Ding, P. Fan, and H. V. Poor, "Impact of non-orthogonal multiple access on the offloading of mobile edge computing," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 375–390, Jan. 2019.

[12] F. Fang, Y. Xu, Z. Ding, C. Shen, M. Peng, and G. K. Karagiannidis, "Optimal resource allocation for delay minimization in NOMA-MEC networks," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7867–7881, Dec. 2020.

[13] B. Liu, C. Liu, and M. Peng, "Resource allocation for energy-efficient MEC in NOMA-enabled massive IoT networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1015–1027, Apr. 2021.

[14] Q.-V. Pham, H. T. Nguyen, Z. Han, and W.-J. Hwang, "Coalitional games for computation offloading in NOMA-enabled multi-access edge computing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1982–1993, Feb. 2020.

[15] A. Kiani and N. Ansari, "Edge computing aware NOMA for 5G networks," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1299–1306, Aug. 2018.

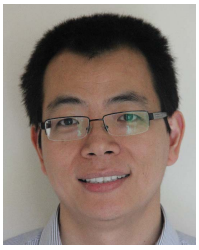
[16] M. Sheng, Y. Dai, J. Liu, N. Cheng, X. Shen, and Q. Yang, "Delay-aware computation offloading in NOMA MEC under differentiated uploading delay," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2813–2826, Apr. 2020.

[17] W. Feng *et al.*, "Hybrid beamforming design and resource allocation for UAV-aided wireless-powered mobile edge computing networks with NOMA," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 11, pp. 3271–3286, Nov. 2021.

[18] I. Budhiraja, N. Kumar, S. Tyagi, and S. Tanwar, "Energy consumption minimization scheme for NOMA-based mobile edge computation networks underlaying UAV," *IEEE Syst. J.*, vol. 15, no. 4, pp. 1–10, Dec. 2021.

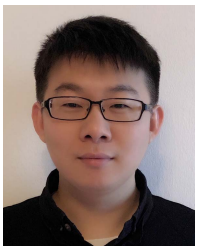
[19] Y. Xu, T. Zhang, D. Yang, Y. Liu, and M. Tao, "Joint resource and trajectory optimization for security in UAV-assisted MEC systems," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 573–588, Jan. 2021.

[20] L. Qian, W. Wu, W. Lu, Y. Wu, B. Lin, and T. Q. S. Quek, "Secrecy-based energy-efficient mobile edge computing via cooperative non-orthogonal multiple access transmission," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4659–4677, Jul. 2021.
[21] B. Li, W. Wu, W. Zhao, and H. Zhang, "Security enhancement with a hybrid cooperative NOMA scheme for MEC system," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2635–2648, Mar. 2021.
[22] B. Li, F. Si, W. Zhao, and H. Zhang, "Wireless powered mobile edge computing with NOMA and user cooperation," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1957–1961, Feb. 2021.
[23] F. Zhou and R. Q. Hu, "Computation efficiency maximization in wireless-powered mobile edge computing networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3170–3184, May 2020.
[24] Z. Ding, J. Xu, O. A. Dobre, and V. Poor, "Joint power and time allocation for NOMA-MEC offloading," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6207–6211, Jun. 2019.
[25] Z. Ding, D. W. K. Ng, R. Schober, and H. V. Poor, "Delay minimization for NOMA-MEC offloading," *IEEE Signal Process. Lett.*, vol. 25, no. 12, pp. 1875–1879, Dec. 2018.
[26] M. Zeng, N.-P. Nguyen, O. A. Dobre, and H. V. Poor, "Delay minimization for NOMA-assisted MEC under power and energy constraints," *IEEE Wireless Commun. Lett.*, vol. 8, no. 6, pp. 1657–1661, Dec. 2019.
[27] I. Altin and M. Akar, "Novel OMA and hybrid NOMA schemes for MEC offloading," in *Proc. IEEE Int. Black Sea Conf. Commun. Netw. (BlackSeaCom)*, Odessa, Ukraine, May 2020, pp. 1–5.
[28] T. Cover and J. Thomas, *Elements of Information Theory*, 6th ed. New York, NY, USA: Wiley, 1991.
[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[30] Z. Ding, R. Schober, and H. V. Poor, "Unveiling the importance of SIC in NOMA systems—Part 1: State of the art and recent findings," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2373–2377, Nov. 2020.
[31] Z. Ding, R. Schober, and H. V. Poor, "Unveiling the importance of SIC in NOMA systems—Part II: New results and future directions," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2378–2382, Nov. 2020.



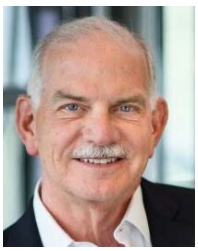
Zhiguo Ding (Fellow, IEEE) received the B.Eng. degree from the Beijing University of Posts and Telecommunications in 2000, and the Ph.D. degree from Imperial College London in 2005.

From July 2005 to April 2018, he was working with Queen's University Belfast, Imperial College, Newcastle University, and Lancaster University. Since April 2018, he has been a Professor of communications with The University of Manchester. From October 2012 to September 2022, he has also been an Academic Visitor with Princeton University. His research interests include 5G networks, game theory, cooperative and energy harvesting networks, and statistical signal processing. He recently received the EU Marie Curie Fellowship 2012–2014, the Top IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY Editor 2017, the IEEE Heinrich Hertz Award 2018, the IEEE Jack Neubauer Memorial Award 2018, the IEEE Best Signal Processing Letter Award 2018, the Friedrich Wilhelm Bessel Research Award 2020, and the IEEE SPCC Technical Recognition Award 2021. He is serving as an Area Editor for the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY and an Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. From 2013 to 2016, he was an Editor of the IEEE WIRELESS COMMUNICATION LETTERS, the IEEE TRANSACTIONS ON COMMUNICATIONS, and the IEEE COMMUNICATION LETTERS. He is a Distinguished Lecturer of IEEE ComSoc, and a Web of Science Highly Cited Researcher in two categories 2021.



Dongfang Xu (Graduate Student Member, IEEE) received the B.S. degree in communication engineering from Shandong University, Jinan, China, in 2014, and the M.S. degree in communication and multimedia engineering from the Friedrich-Alexander University of Erlangen-Nuremberg (FAU), Germany, in 2017, where he is currently pursuing the Ph.D. degree in electrical engineering.

Robert Schober (Fellow, IEEE) received the Diploma (Univ.) and Ph.D. degrees in electrical engineering from the Friedrich-Alexander University of Erlangen-Nuremberg (FAU), Germany, in 1997 and 2000, respectively. From 2002 to 2011, he was a Professor and a Canada Research Chair at The University of British Columbia (UBC), Vancouver, BC, Canada. Since January 2012, he has been an Alexander von Humboldt Professor and the Chair for Digital Communication at FAU. His research interests include communication theory, wireless communications, and statistical signal processing. He is a fellow of the Canadian Academy of Engineering, a fellow of the Engineering Institute of Canada, and a member of the German National Academy of Science and Engineering. He received several awards for his work, including the 2002 Heinz Maier Leibnitz Award of the German Science Foundation (DFG), the 2004 Innovations Award of the Vodafone Foundation for Research in Mobile Communications, the 2006 UBC Killam Research Prize, the 2007 Wilhelm Friedrich Bessel Research Award of the Alexander von Humboldt Foundation, the 2008 Charles McDowell Award for Excellence in Research from UBC, the 2011 Alexander von Humboldt Professorship, the 2012 NSERC E.W.R. Stacie Fellowship, and the 2017 Wireless Communications Recognition Award by the IEEE Wireless Communications Technical Committee. Since 2017, he has been listed as a Highly Cited Researcher by the Web of Science. From 2012 to 2015, he served as the Editor-in-Chief for the IEEE TRANSACTIONS ON COMMUNICATIONS. He currently serves as a member of the Editorial Board for the PROCEEDINGS OF THE IEEE and as the VP Publications for the IEEE Communication Society (ComSoc).



H. Vincent Poor (Life Fellow, IEEE) 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 at Princeton University, where he is currently the Michael Henry Strater University Professor. From 2006 to 2016, he served as the Dean of the 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 include information theory, machine learning and network science, and their applications in wireless networks, energy systems, and related fields. Among his publications in these areas is the forthcoming book *Machine Learning and Wireless Communications* (Cambridge University Press). Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and is a foreign member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. He received the IEEE Alexander Graham Bell Medal in 2017.