Joint Rate Control and Power Allocation for Non-Orthogonal Multiple Access Systems

Wei Bao, *Member, IEEE*, He Chen, *Member, IEEE*, Yonghui Li, *Senior Member, IEEE*, and Branka Vucetic, *Fellow, IEEE*

Abstract—This paper investigates the optimal resource allocation of a downlink non-orthogonal multiple access (NOMA) system consisting of one base station and multiple users. Unlike existing short-term NOMA designs that focused on the resource allocation for only the current transmission timeslot, we aim to maximize a long-term network utility by jointly optimizing the data rate control at the network layer and the power allocation among multiple users at the physical layer, subject to practical constraints on both the short-term and long-term power consumptions. To solve this problem, we leverage the recently developed Lyapunov optimization framework to convert the original long-term optimization problem into a series of online rate control and power allocation problems in each timeslot. The power allocation problem, however, is shown to be non-convex in nature and thus cannot be solved with a standard method. However, we explore two structures of the optimal solution and develop a dynamic programming-based power allocation algorithm, which can derive a globally optimal solution, with a polynomial computational complexity. Extensive simulation results are provided to evaluate the performance of the proposed joint rate control and power allocation framework for NOMA systems, which demonstrate that the proposed NOMA design can significantly outperform multiple benchmark schemes, including orthogonal multiple access schemes with optimal power allocation and NOMA schemes with non-optimal power allocation, in terms of average throughput and data delay.

Index Terms—Non-orthogonal multiple access, rate control, power allocation, Lyapunov optimization.

I. Introduction

DUE to the explosive traffic growth and the fast proliferation of Internet of Things (IoT), the fifth generation (5G) of cellular networks are expected to face unprecedented challenges including 1000-fold increase in system capacity, improved spectral efficiency, and massive connectivity with diverse service requirements [1], [2]. In this context, non-orthogonal multiple access (NOMA), although not completely new to the wireless industry and research community [3], has been regarded as a promising radio access technology for the 5G wireless communication systems [4]–[7], due to its

unique capability of achieving a higher spectral efficiency and supporting a large number of concurrent transmissions over the same communication resource. In fact, multiuser superposition transmission (MUST), a two-user downlink scenario of NOMA, has been investigated for the third generation partnership project long-term evolution advanced (3GPP-LTEA) networks [8].

The current fourth generation (4G) of cellular communication systems and previous generations primarily adopted orthogonal multiple access (OMA) technologies, such as frequency-division multiple access (FDMA), time-division multiple access (TDMA), code-division multiple access (CDMA), and orthogonal frequency-division multiple access (OFDMA). In these OMA schemes, the resources are first split into orthogonal resource blocks in frequency/time/code domain and each resource block is then assigned to one user exclusively. As such, the inter-user interference can be avoided and the information of each user can be recovered at a low complexity. However, due to the orthogonal resource allocation, the maximum number of served users is limited by the total number of available resource blocks. In this sense, the OMA techniques are inadequate to support the massive connectivity requirement of 5G wireless systems. Another main problem of the OMA schemes is their relatively low spectral efficiency, especially for the resource blocks assigned to the users with poor instantaneous channel conditions. This issue can be effectively addressed by applying user selection schemes, where the users with strong channel conditions are selected out to transmit over the limited number resources. However, this may pose a serious user fairness problem as the users near to the cell edge could have much fewer transmission opportunities than those near to the cell center.

Unlike OMA with orthogonal resource allocation, NOMA advocates the usage of the power domain to multiplex signal streams from multiple users together and serve them simultaneously using the same frequency/time/code resource block. At the transmitter, NOMA adopts the superposition coding [9] to superimpose the signals of multiple users together by splitting them in the power domain. At the receiver side, successive interference cancellation (SIC) [10] is implemented to separate multiplexed users' signals. In this case, each user can access to all resource blocks such that those resources that are solely assigned to users with poor channel quality in OMA can still be accessed by other users with good channel conditions in NOMA, which enables NOMA to achieve a higher spectral efficiency than OMA [11]. Apart from this, NOMA is capable of realizing an improved tradeoff between system throughput and user fairness than OMA. This can be

achieved by allocating less power to users with better channel conditions and more power to users with worse channel conditions, which is totally opposite to the conventional water-filling power allocation scheme widely used in OMA.

A. Related Work

The idea of using NOMA as a potential candidate for 5G multiple access technology was first proposed in [12] and [13] by NTT DoCoMo, as a part of the mobile and wireless communications enablers for the twenty-twenty information society (METIS) projects. The system-level performance of NOMA was evaluated by simulation in [14] by taking into consideration many practical factors, including multiuser power allocation, signaling overhead, SIC error propagation and high user mobility. A test-bed of two-user downlink NOMA system was developed and evaluated in [15], where a NOMA scheme with each user occupying the whole bandwidth of 5.4 MHz was compared to an OMA scheme with each user using the transmission bandwidth of 2.7 MHz. Experimental results demonstrated that the NOMA scheme significantly outperforms its OMA counterpart in terms of both aggregate and individual user's throughput.

Since NOMA uses the power domain to realize multiple access and implements SIC to perform user decoding, the power allocation to the data flows of different users plays an important role in determining the performance of NOMA systems. Specifically, the total power of the transmitter should be allocated in a proper way so that the signals for the users with worse channel conditions can be successfully decoded and subsequently subtracted from the received signal of those users with better channel conditions. Early efforts on NOMA have mostly adopted the fixed power allocation scheme, in which the power allocation coefficients are predetermined and are not affected by the instantaneous channel conditions [16]–[20]. The fixed power allocation strategy is favorable in terms of low implementation complexity, however, the system performance can be disadvantaged if the power allocation coefficients are not set appropriately. Inspired by the key idea of cognitive radio, another power allocation scheme was proposed and analyzed in [21], in which NOMA is treated as a special case of cognitive radio networks. In this scheme, the user with a weaker channel condition is regarded as a primary user, who has a higher priority to be served. That is, the available power is first allocated to fulfill the quality of service of the primary user, while the secondary user with a better channel condition will be opportunistically served using the remaining amount of power. In this case, the overall system spectral efficiency could be limited by the fact that the strong user's performance is highly affected by the weak user's channel quality. Recently, Yang *et al.* proposed a general power allocation scheme to protect the QoS of both users, in which the power allocation factors are dynamically adjusted according to the instantaneous channel status to ensure that the rates of two users in NOMA are both larger than those in OMA [22], [23].

The aforementioned power allocation strategies were primarily designed for two-user NOMA scenarios. In general, more than two users can be multiplexed on the same resources, where the power allocation problem becomes more complicated. In [24], the dynamic power allocation scheme for a multi-user NOMA system was addressed with max-min user fairness for the cases with and without instantaneous channel state information (CSI). More specifically, the objective was set to maximize the minimum achievable user rate when the instantaneous CSI is available, and to minimize the maximum user outage probability when only average CSI is known. For the case without instantaneous CSI, a sum power minimization problem was resolved in [25] subject to outage probabilistic constraints and the optimal decoding order. The work of [26] provided an optimal power allocation solution to maximize the weighted sum rate of all users subject to a total power constraint.

When NOMA is applied in a multi-carrier system, the power allocation problem will be upgraded to a joint power and channel allocation problem. There have been some great efforts in this research line [27]–[33]. The joint power and channel allocation problem of multi-carrier NOMA systems was formally formulated with both maximum weighted sum rate and sum-rate utilities in [27] and [28], in which the problem was proved to be NP hard and a new algorithm inspired by Lagrangian duality and dynamic programming was proposed to achieve a near-optimal solution. In [29] and [30], the joint power and channel allocation problem was resolved by decoupling it as a many-to-many matching game with externalities and a geometric programming subproblem. Apart from the maximization of the sum rate or weighted sum rate, the total transmit power minimization problems were addressed subject to the predefined QoS of individual users in [31] and [32]. Very recently, Sun *et al.* studied a full-duplex multi-carrier NOMA system, wherein the monotonic optimization was adopted to design an optimal joint power and subcarrier allocation policy such that the weighted sum system throughput is maximized.

B. Motivation and Contributions

We notice that most existing dynamic resource allocation strategies for NOMA systems have mainly focused on short-term (e.g., "one-snapshot") designs. Specifically, the system performance is normally optimized subject to a short-term peak power constraint for a single transmission block, overlooking the practical long-term power constraint. These short-term designs may lead to inferior system performance in a long-term perspective since the power resources are forced to be used even when the channel condition is not good enough in some transmission blocks. Furthermore, in practice, the amount of data that can be transmitted at the physical layer is highly influenced by the rate control at the network layer. The limited transmit power should not be wasted to those users who have little data to send/receive. In this sense, to achieve a higher long-term system throughput with limited resources, it is desirable to jointly design the rate control at the network layer and the resource allocation at the physical layer for NOMA systems. However, to our best knowledge, this important problem has not been considered in the open literature.

Motivated by this gap, in this paper we investigate the joint rate control and power allocation of a downlink NOMA

system with one base station (BS) communicating to multiple users. Different from existing short-term resource allocation policies, we aim to maximize the long-term network utility, defined as a sum of concave functions of average data rates of all users, subject to a peak power constraint and a long-term average power constraint in the physical layer, as well as a peak rate constraint in the network layer. In the proposed design, we implement data queues at the BS to characterize users' incoming data flows from the Internet and outgoing data flows to the wireless channel. Note that it is desirable to jointly consider the data rate control and power optimization problems due to the fact that the amount of data delivered to each user in a certain timeslot is limited by not only the instantaneous channel capacity but also the amount of data available in the queue. Furthermore, the considered short-term peak power constraint and long-term average power constraint make the rate control and power allocation of different users tangled not only in each timeslot but also across various timeslots in the long term. In this sense, developing an online algorithm that can maximize the long-term network utility through optimizing the coupled data rate control and power allocation in each timeslot subject to both short-term and long-term power constraints is a challenging task.

The main contributions of this paper are summarized as follows:

- We, for the first time, develop a systematic framework for the joint rate control and power allocation in NOMA systems. The formulated long-term optimization problem is fundamentally different from those short-term designs in the open literature since it is subject to *long-term* average and *short-term* peak power constraints, the network-layer peak rate constraints, and the queue stability constraints. As such, it cannot be resolved by the existing approaches. Motivated by this, we leverage the recently-developed Lyapunov optimization approach [34], [35], which enables us to achieve the asymptotic optimality of the formulated problem. This is realized by converting the original long-term optimization problem to a series of online queue- and channel-aware optimization problems to be resolved at each single timeslot. Here, "online" means that the BS only needs to know the queue states and channel states at the current timeslot, without predicting anything in the future or knowing the statistics of the channel states.
- The online queue- and channel-aware optimization problem is then decomposed into rate control and power allocation problems in each timeslot. However, the power allocation problem is non-convex in nature and cannot be solved with a standard method. Instead, we explore two important structures in solving the problem. Referred to as the *finite-point structure*, we show that the optimality only possibly happens at a finite set of candidate points, which substantially decreases the searching space. Referred to as the *incremental structure*, we derive the recursion of the original optimization problem and its subproblems, leading to the desired Bellman equation. As a consequence, we propose a new dynamic programming based power allocation (DPPA) approach to
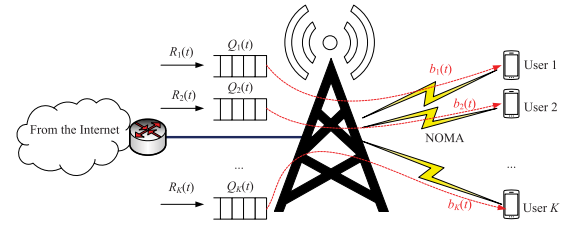

Fig. 1.   System model of the considered NOMA system.

derive a globally optimal solution, within a polynomial computational complexity.

- We evaluate the performance of our proposed framework for joint rate control and power allocation for NOMA systems via simulation and show that NOMA can greatly improve the network performance in both data rate and delay compared with four benchmark schemes, including OMA and non-optimal NOMA schemes, under a variety of system settings.

C. Organization

The rest of the paper is organized as follows. The system model and problem formulation are described in Section II, where the Lyapunov optimization approach is applied to convert the original long-term optimization problem into a series of rate control and power allocation problems to be optimized in each timeslot. In Section III, we elaborate the two special structures of the power allocation problems and develop a new dynamic programming based power allocation algorithm, which is guaranteed to achieve its global optimality. Simulation results and the associated discussions are presented in Section IV, and the conclusions of the paper are drawn in Section V.

II. System Model and Lyapunov Optimization

A. System Model

As shown in Fig. 1, we consider the downlink data transmission from one BS to K users using NOMA [16], [20], [24], [25], [28], [29]. All the nodes are equipped with single antenna and work in a half-duplex mode. The system operates on slotted time $t \in \{0, 1, 2, \ldots\}$. At timeslot t, $R_i(t)$ bits of data for user i arrive at the BS from the Internet. These data are firstly buffered at queue i and then forwarded to user i via a wireless channel. Let $Q_i(t)$ denote the amount of data buffered at queue i at time t, i.e., queue backlog of the user i. We assume that the queue is large enough so that no data will be dropped. Let $b_i(t)$ denote the amount of data that can be delivered to the user i at t (i.e., data transmission capability offered by the underlying wireless channel). $R_i(t)$ and $b_i(t)$ are variables to be designed by the system. $R_i(t)$ indicates how many data should be sent to the BS in the view of the network layer, and $b_i(t)$ indicates how many data should be sent from the BS to user i via the wireless link in the view of the physical layer. Since the data rate of one user cannot be arbitrarily large in the network, we consider that $R_i(t)$ is limited to $R_{\max}$.[1]

---

[1]This value is imposed by an upper-layer protocol (e.g., TCP) for each user. This is because the upper layer protocol usually maintains a buffer to send its data, and maximum possible data can be sent in each timeslot is limited by the buffer size. This assumption is widely adopted in the literature that focuses on network flow control, such as [34, Sec. 5.2.4] and [35, Sec. 5.2].

$b_i(t)$ is also limited by finite transmission power, which will be elaborated shortly. Mathematically, we can write the following expression to characterize the evolution of each user's queue backlog

$$Q_i(t+1) = [Q_i(t) - b_i(t)]^+ + R_i(t), \quad (1)$$

where $[\cdot]^+$ means $\max[\cdot, 0]$.

To improve the spectral efficiency, in this paper we consider that the BS adopts the NOMA method to transmit data to users in the downlink. As the first effort towards the long-term resource allocation designs for NOMA, here we concentrate on a single-carrier NOMA system such that the data transmissions to all users work on the same frequency band. The proposed long-term resource allocation framework can be extended to a multi-carrier NOMA system, which we would like to consider as future work. In addition, it is assumed that channels from the BS to the users experience quasi-static and frequency flat fading such that the channel gains remain constant during each timeslot but change independently from one timeslot to another. We use $\widehat{g}_i(t)$ to denote the channel gain from the BS to user i during timeslot t. $\widehat{g}_i(t)$ are independent in different timeslots, but $\widehat{g}_i(t)$ and $\widehat{g}_j(t)$, $i \neq j$, can be dependent in one timeslot. As explained later, in our framework we do not need to know the probability distribution of $\widehat{g}_j(t)$. Let $p_i(t)$ denote the power allocated to transmit the data of user i at t. Let $\widehat{\mathbf{g}}(t) \triangleq (\widehat{g}_1(t), \widehat{g}_2(t), \dots, \widehat{g}_K(t))$ and $\mathbf{p}(t) \triangleq (p_1(t), p_2(t), \dots, p_K(t))$. $\mathbf{p}(t)$ is a power allocation variable that determines how the power is allocated to multiple users at t. As we will see shortly, the values of $b_i(t)$ depend on channel gains and power allocations at t.

We define $g_{s(1)}(t), g_{s(2)}(t), \dots, g_{s(K)}(t)$ as sorted $\widehat{g}_1(t), \dots, \widehat{g}_K(t)$ in a descending order, where $s(i)$ indicates the index of the user with the ith largest channel gain, i.e., subscript of the user before sorting. According to the principle of NOMA with superposition coding and SIC, we can compute the amount of data that can be sent to user $s(i)$ (i.e., $b_{s(i)}(t)$) based on the channel gains and power allocation given below [24], [28]:

$$\begin{cases} b_{s(1)}(t) = W\tau \log\left(1 + \frac{p_{s(1)}(t)g_{s(1)}(t)}{\eta}\right), \\ b_{s(2)}(t) = W\tau \log\left(1 + \frac{p_{s(2)}(t)g_{s(2)}(t)}{p_{s(1)}(t)g_{s(2)}(t)+\eta}\right), \\ \dots \\ b_{s(K)}(t) = W\tau \log\left(1 + \frac{p_{s(K)}(t)g_{s(K)}(t)}{\sum_{j=1}^{K-1} p_{s(j)}(t)g_{s(K)}(t)+\eta}\right), \end{cases} \quad (2)$$

where W is the radio frequency bandwidth, τ is the duration of one timeslot, and η is the noise level. W, τ, and η are predetermined.

In reality, the BS transmission power can be constrained by two practical limitations. One could be imposed by the requirement for power savings, which limits the long-term average power consumption at the BS. The other may come from the regulations and rules (e.g., Federal Communication Commission (FCC)), which restricts the short-term transmission power at the BS. Motivated by this, we consider that the NOMA system is subject to a long-term average power

constraint,[2]

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[\sum_{i=1}^{K} p_i(t)\right] \leq P_{\text{mean}}, \quad (3)$$

as well as a one-timeslot peak power constraint,

$$\sum_{i=1}^{K} p_i(t) \leq P_{\text{max}}, \quad \forall t, \quad (4)$$

where P_{mean} and P_{max} are the maximum allowed long-term average power usage, and maximum allowed instantaneous power usage in each timeslot, respectively. Please note that the existing resource allocation designs for NOMA normally considered the short-term power constraint, but overlooked the long-term power constraint.

B. Problem Statement

To proceed, we first define the long-term network utility as follows

$$\mathcal{U} = \sum_{i=1}^{K} U_i\left(\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[R_i(t)]\right), \quad (5)$$

where $U_i(\cdot)$ is the utility function of user i. It is an arbitrary concave non-decreasing function. User i's utility is a function of its long-term average data rate, and the overall network utility is the sum of the individual user utilities.

The aim of this paper is to optimize the long-term network utility, by designing the data rate control in the network layer, $\mathbf{R}(t)$, and the power allocation in the physical layer, $\mathbf{p}(t)$. The problem is formally stated as the following Problem Origin (Problem PO):

$$\max_{\mathbf{R}(t),\mathbf{p}(t),\forall t} \sum_{i=1}^{K} U_i\left(\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[R_i(t)]\right), \quad (6a)$$

$$\text{subject to } \sum_{i=1}^{K} p_i(t) \leq P_{\text{max}}, \quad \forall t, \quad (6b)$$

$$p_i(t) \geq 0, \quad \forall i, t, \quad (6c)$$

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[\sum_{i=1}^{K} p_i(t)\right] \leq P_{\text{mean}}, \quad (6d)$$

$$0 \leq R_i(t) \leq R_{\text{max}}, \quad \forall i, t, \quad (6e)$$

$$\lim_{t\to\infty} \frac{\mathbb{E}[Q_i(t)]}{t} = 0, \quad (6f)$$

where $\mathbf{R}(t) \triangleq (R_1(t), R_2(t), \dots, R_K(t))$. The last constraint $\lim_{t\to\infty} \frac{\mathbb{E}[Q_i(t)]}{t} = 0$ means that $Q_i(t)$ is *mean rate stable*. Its physical meaning is that the queue will not grow to infinity after sufficiently long period. Note that due to the considered queueing model given in (1) and the long-term average power constraint given in (3), the rate control and

---

[2]The long-term average power constraint at the BS could be imposed by the mobile operator to control its electricity expense. It could also be imposed by the government agency to limit the energy consumption of BS to reduce its carbon emission on average. It is worth mentioning that this type of average power constraint has been commonly considered in existing literature that focused on the optimization of long-term performance of wireless systems, see e.g., [36], [37].

power allocation of all users across different timeslots are actually tangled together. As such, the optimal solution to the formulated optimization problem (6) cannot be achieved by simply maximizing the data rate in each timeslot.

C. Lyapunov Optimization

The Problem PO perfectly matches the Lyapunov optimization framework [34], [35]. This is because we aim to maximize a long-term utility under a long-term constraint as well as queue stability constraints. It can be shown that the long-term power constraint (6d) can be converted to a virtual queue stable constraint [34], [35]. We define a virtual queue $Z(t)$, with $Z(0) = 0$ and update equation

$$Z(t + 1) = \left[Z(t) + \sum_{i=1}^{K} p_i(t) - P_{\text{mean}} \right]^+. \tag{7}$$

With reference to [35, Sec. 4.4], (6d) can be guaranteed if virtual queue $Z(t)$ is mean rate stable. As such, we replace condition (6d) by "$Z(t)$ is mean rate stable". In fact, $Z(t)$ indicates the accumulated power debt at t: On average, the amount of P_{mean} power can be consumed in each timeslot. Using power more than P_{mean} generates power debt, and using power less than P_{mean} repays the power debt. The power debt must be maintained finite over infinite timeslots, so that the long-term power constraint is satisfied.

Subsequently, we study the queue evolution of the system. Let $\mathbf{Q}(t) \triangleq (Q_1(t), Q_2(t), \ldots, Q_K(t))$, and $\Theta(t) = (\mathbf{Q}(t), Z(t))$. The Lyapunov function $\mathbf{L}(\Theta(t))$ then becomes [34], [35]

$$\mathbf{L}(\Theta(t)) \triangleq \frac{1}{2} \left[\sum_{i=1}^{K} Q_i^2(t) + Z^2(t) \right]. \tag{8}$$

The Lyapunov drift is then defined as

$$\Delta(\Theta(t)) \triangleq \mathbb{E} \left[\mathbf{L}(\Theta(t + 1)) - \mathbf{L}(\Theta(t)) | \Theta(t) \right]. \tag{9}$$

Since we aim to maximize a long-term utility as well as to guarantee the queue stability, we focus on the following drift minus utility (drift plus penalty [34], [35])

$$\text{DPP}(\Theta(t)) = \Delta(\Theta(t)) - V \mathbb{E} \left[\sum_{i=1}^{K} U_i(R_i(t)) \Big| \Theta(t) \right], \tag{10}$$

where V is a tunable weight, representing the relative importance of "utility maximization" compared with "queue stability". Then $\text{DPP}(\Theta(t))$ can be further bounded in the following way [35, Sec. 3.1.2]

$\text{DPP}(\Theta(t))$
$$\leq \mathbb{E} \left[\sum_{i=1}^{K} \frac{R_i^2(t) + b_i^2(t)}{2} \Big| \Theta(t) \right] + \mathbb{E} \left[\sum_{i=1}^{K} Q_i(t) R_i(t) \Big| \Theta(t) \right]$$
$$- \mathbb{E} \left[\sum_{i=1}^{K} Q_i(t) b_i(t) \Big| \Theta(t) \right]$$
$$+ \mathbb{E} \left[\frac{\left(\sum_{i=1}^{K} p_i(t) - P_{\text{mean}} \right)^2}{2} \Big| \Theta(t) \right]$$

$$+ \mathbb{E} \left[Z(t) \left(\sum_{i=1}^{K} p_i(t) - P_{\text{mean}} \right) \Big| \Theta(t) \right]$$
$$- V \mathbb{E} \left[\sum_{i=1}^{K} U_i(R_i(t)) \Big| \Theta(t) \right]$$
$$\leq B + \mathbb{E} \left[\sum_{i=1}^{K} Q_i(t) R_i(t) \Big| \Theta(t) \right] - \mathbb{E} \left[\sum_{i=1}^{K} Q_i(t) b_i(t) \Big| \Theta(t) \right]$$
$$+ \mathbb{E} \left[Z(t) \left(\sum_{i=1}^{K} p_i(t) - P_{\text{mean}} \right) \Big| \Theta(t) \right]$$
$$- V \mathbb{E} \left[\sum_{i=1}^{K} U_i(R_i(t)) \Big| \Theta(t) \right], \tag{11}$$

where B is an upper bound of $\mathbb{E} \left[\sum_{i=1}^{K} \frac{R_i^2(t) + b_i^2(t)}{2} \Big| \Theta(t) \right] +$ $\mathbb{E} \left[\frac{\left(\sum_{i=1}^{K} p_i(t) - P_{\text{mean}} \right)^2}{2} \Big| \Theta(t) \right]$. Please note that this value is bounded since $R_i(t)$, $b_i(t)$, and $p_i(t)$ are all bounded values.

Following the Lyapunov optimization approach, we need to "opportunistically" minimize the drift minus utility at each timeslot, denoted by $\text{obj}(t)$, i.e., the following term is minimized at each timeslot [34], [35]

$\text{obj}(t)$
$$= \sum_{i=1}^{K} Q_i(t) R_i(t) - \sum_{i=1}^{K} Q_i(t) b_i(\mathbf{p}(t), \widehat{\mathbf{g}}(t))$$
$$+ Z(t) \left(\sum_{i=1}^{K} p_i(t) - P_{\text{mean}} \right) - V \sum_{i=1}^{K} U_i(R_i(t)) \tag{12}$$
$$= \sum_{i=1}^{K} [Q_i(t) R_i(t) - V U_i(R_i(t))]$$
$$- \left[\sum_{i=1}^{K} Q_i(t) b_i(\mathbf{p}(t), \widehat{\mathbf{g}}(t)) - Z(t) \left(\sum_{i=1}^{K} p_i(t) - P_{\text{mean}} \right) \right]. \tag{13}$$

Please note that $b_i(t)$ is a function of $\mathbf{p}(t)$ and $\widehat{\mathbf{g}}(t)$ as shown in (2), so that we use $b_i(\mathbf{p}(t), \widehat{\mathbf{g}}(t))$ to represent $b_i(t)$ in the above equation. As a consequence, we arrive at the following single-timeslot optimization, which is referred to as Problem Single-Timeslot (Problem PST)

$$\min_{\mathbf{p}(t), \mathbf{R}(t)} \quad \text{obj}(t), \tag{14a}$$
$$\text{subject to} \quad \sum_{i=1}^{K} p_i(t) \leq P_{\max}, \tag{14b}$$
$$p_i(t) \geq 0, \quad \forall i, \tag{14c}$$
$$0 \leq R_i(t) \leq R_{\max}, \quad \forall i. \tag{14d}$$

According to (13)–(14d), we notice that in order to solve the Problem PST, the BS only needs to know the queue states $Q_i(t), \forall i$, virtual queue state $Z(t)$, and channel gains $\widehat{g}_i(t), \forall i$, in the current timeslot t. Therefore, the solution to Problem PST is an online queue- and channel-aware solution.

D. Optimality of Lyapunov Optimization

Our original aim is to solve the long-term optimization Problem PO. By employing the Lyapunov optimization approach, we convert the problem into the online single-timeslot optimization Problem PST. If the single-timeslot optimization is solved optimally at each timeslot, according to [34, Sec. 5.4] and [35, Sec. 5.1], the optimal solution to the original Problem PO can then be achieved asymptotically, following an $[O(V), O(\frac{1}{V})]$ tradeoff between the queue backlogs and the achieved utility given below

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{K} \mathbb{E}(Q_i(t)) \leq \frac{B}{\mu_{\text{sym}}} + O(V), \quad (15)$$

$$\sum_{i=1}^{K} U_i \left(\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[R_i(t)] \right) \geq U^* - O\left(\frac{1}{V}\right), \quad (16)$$

where U^* represents the maximum possible utility of Problem PO, μ_{sym} is a constant (predetermined by the system) representing the largest possible time average data rate that is simultaneously supportable to all users,[3] and $O(\cdot)$ represents the big O notation.

E. Decomposition of Problem PST

Since the objective function of Problem PST is in a sum form and the problem contains non-coupled constraints, Problem PST can be decomposed into the following two subproblems. First, for each user i, we need to optimize

$$\max_{R_i(t)} \quad V U_i(R_i(t)) - Q_i(t) R_i(t), \quad (17a)$$

$$\text{subject to} \quad 0 \leq R_i(t) \leq R_{\max}. \quad (17b)$$

The above problem is a rate control problem for user i. It is referred to as Problem RC-i in the rest of this paper. Second, the following problem must also be solved

$$\max_{\mathbf{p}(t)} \quad \sum_{i=1}^{K} Q_i(t) b_i(\mathbf{p}(t), \widehat{\mathbf{g}}(t)) - Z(t) \sum_{i=1}^{K} p_i(t), \quad (18a)$$

$$\text{subject to} \quad \sum_{j=1}^{K} p_j(t) \leq P_{\max}, \quad (18b)$$

$$p_i(t) \geq 0, \quad \forall i. \quad (18c)$$

[3] μ_{sym} represents the largest possible time average data rate that is simultaneously supportable to all user [34, Definition 3.7 and Sec. 5.1.2]. Its physical meaning is explained as follows: Suppose that the mean arrival rates of the K users $(R_1(t), R_2(t), \ldots, R_K(t))$ are equal to $(\mu_1, \mu_2, \ldots, \mu_K)$. Then, $(\mu_1, \mu_2, \ldots, \mu_K)$ is in the network-layer capacity region if these arrival rates can be stably supported by the network (i.e., all queues are stable), considering all possible strategies for choosing the control variables to affect scheduling and resource allocation. $(\mu_1, \mu_2, \ldots, \mu_K)$ is not in the network-layer capacity region if it is not possible to find any strategy to support these arrival rates. μ_{sym} is the value such that $(\mu_{\text{sym}}, \mu_{\text{sym}}, \ldots, \mu_{\text{sym}})$ is in the capacity region but $(\mu_{\text{sym}} + \epsilon, \mu_{\text{sym}} + \epsilon, \ldots, \mu_{\text{sym}} + \epsilon)$ is out of the capacity region for an arbitrary small ϵ. For an arbitrary system, μ_{sym} exists and can be regarded as a predetermined constant. Please note that μ_{sym} is usually a hidden value when we design the system. In this paper, we do not need to know μ_{sym} to design rate control or power allocation. It is only used to show the $[O(V), O(\frac{1}{V})]$ tradeoff between the queue backlogs and the achieved utility.

This is a power allocation problem for all users. It is referred to as Problem PA in the rest of the paper. Problem PST is solved optimally if and only if Problems RC-i and Problem PA are all solved optimally.

The solution to Problem RC-i is straightforward. It is a single-variable optimization problem and the objective function is concave. The optimal $R_i(t)$ is simply equal to the solution to $V \frac{\partial U_i(t)}{\partial R_i(t)} = Q_i(t)$ (if it is grater than $R_{\max}$ or smaller than 0, then the optimal $R_i(t)$ equals to $R_{\max}$ or 0 respectively).

However, we can easily verify that Problem PA is a non-convex optimization problem, which cannot be solved through a standard method. For a general non-convex optimization problem, it is not guaranteed that there is an algorithm that can find a globally optimal solution within a polynomial computational complexity. Nevertheless, in Section III, we manage to develop two important structures for solving the problem, so that a globally optimal solution to Problem PA can be calculated within a polynomial computational complexity, through the proposed Dynamic Programming based Power Allocation (DPPA) algorithm.

III. DYNAMIC PROGRAMMING BASED POWER ALLOCATION

In this section, we aim to solve the single-timeslot power allocation problem, i.e., Problem PA. Our main contribution is to propose a new dynamic programming approach to derive its globally optimal solution. To this end, we first reformulate Problem PA for presentation convenience. Then, we explore two important properties of the optimization problem, which will be employed to construct the Bellman formula [38, Ch. 15]. Finally, the Dynamic Programming based Power Allocation (DPPA) algorithm is proposed to find the optimal solution to Problem PA.

A. Reformulation of Problem PA

We first reformulate Problem PA to facilitate the presentation. First, problem PA is a single-timeslot optimization problem, which does not depend on other timeslots, so that we ignore the notation t throughout this section. Second, without loss of generality, we assume that the channel gains follow $g_1 \geq g_2 \geq \ldots \geq g_K$. Let $\mathbf{g} \triangleq (g_1, g_2, \ldots, g_K)$. Please note that we do not make any specific assumptions on $Q_1, Q_2, \ldots, Q_K$, and Z so that they could be arbitrary non-negative values. Also, without loss of generality, W and τ are normalized to 1 throughout this section. As a consequence, Problem PA is rewritten as

$$\max_{\mathbf{p}} \quad Q_1 \log\left(1 + \frac{p_1 g_1}{\eta}\right) + Q_2 \log\left(1 + \frac{p_2 g_2}{p_1 g_2 + \eta}\right) + \ldots$$
$$+ Q_K \log\left(1 + \frac{p_K g_K}{\sum_{j=1}^{K-1} p_j g_K + \eta}\right) - Z\left(\sum_{j=1}^{K} p_j\right), \quad (19a)$$

$$\text{s.t.} \quad \sum_{j=1}^{K} p_j \leq P_{\max}, \quad (19b)$$

$$p_k \geq 0, \quad \forall k. \quad (19c)$$

The physical meaning of the above Problem PA is interpreted as follows. First, the objective function (19a) comprises a weighted sum of physical layer data rates of individual users. The weights are $Q_1, Q_2, \ldots, Q_K$ (i.e., the queue backlogs), which means that if more outstanding data are stored at queue i, then data transmission to user i has a higher priority, and this is why its data rate is multiplied by the queue backlog Q_i. The objective function (19a) also includes a penalty term $Z \cdot \sum_{j=1}^{K} p_j$, the product of Z and total power consumption of the current timeslot. Recall from (7) that Z indicates the accumulated power debt. A larger Z value will lead to a higher penalty in power consumption, so that less power will be consumed in the current timeslot, and thus the power debt could be kept finite, i.e., the long-term power constraint is satisfied. It is worth mentioning that the weighted sum rate maximization problem has been studied in [27]–[29]. However, their methods cannot address our power allocation problem due to the different objective function introduced by the considered long-term average power constraint.

In what follows, we introduce two important properties of Problem PA, which will enable to find a globally optimal solution.

B. Finite-Point Structure

In this subsection, we show the finite-point structure of Problem PA, which will greatly decrease the searching space for the optimal solution. Let F denote the objective function (19a).

Lemma 1 (Finite-Point Structure): If the optimality of Problem PA is achieved, $\forall k \leq K$, the value of $p_1 + p_2 + \ldots + p_k$ can only be one from the following values in $\mathcal{E}$, where $\mathcal{E} = \left\{ 0; \frac{\eta \, (g_j Q_j - g_i Q_i)}{g_i g_j (Q_i - Q_j)}, \forall i, j, 1 \leq i < j \leq K; \frac{Q_i}{Z} - \frac{\eta}{g_i}, \forall i, 1 \leq i \leq K; P_{\max} \right\}$.

Proof: To proceed, we first calculate the first-order derivative of the objective function given by

$$
\begin{cases}
\frac{\partial F}{\partial p_1} = \frac{g_1 Q_1}{p_1 g_1 + \eta} - \frac{g_2 Q_2}{p_1 g_2 + \eta} + \frac{g_2 Q_2}{(p_1 + p_2) g_2 + \eta} - \frac{g_3 Q_3}{(p_1 + p_2) g_3 + \eta} + \\
\quad \ldots + \frac{g_{K-1} Q_{K-1}}{(p_1 + \ldots + p_{K-1}) g_{K-1} + \eta} - \frac{g_K Q_K}{(p_1 + \ldots + p_{K-1}) g_K + \eta} \\
\quad + \frac{g_K Q_K}{(p_1 + \ldots + p_K) g_K + \eta} - Z, \\
\frac{\partial F}{\partial p_2} = \frac{g_2 Q_2}{(p_1 + p_2) g_2 + \eta} - \frac{g_3 Q_3}{(p_1 + p_2) g_3 + \eta} + \ldots \\
\quad + \frac{g_{K-1} Q_{K-1}}{(p_1 + \ldots + p_{K-1}) g_{K-1} + \eta} - \frac{g_K Q_K}{(p_1 + \ldots + p_{K-1}) g_K + \eta} \\
\quad + \frac{g_K Q_K}{(p_1 + \ldots + p_K) g_K + \eta} - Z, \\
\ldots \\
\frac{\partial F}{\partial p_{K-1}} = \frac{g_{K-1} Q_{K-1}}{(p_1 + \ldots + p_{K-1}) g_{K-1} + \eta} - \frac{g_K Q_K}{(p_1 + \ldots + p_{K-1}) g_K + \eta} \\
\quad + \frac{g_K Q_K}{(p_1 + \ldots + p_K) g_K + \eta} - Z, \\
\frac{\partial F}{\partial p_K} = \frac{g_K Q_K}{(p_1 + \ldots + p_K) g_K + \eta} - Z.
\end{cases}
\tag{20}
$$

If the optimality is achieved, the following Karush-Kuhn-Tucker (KKT) condition must be satisfied.[4]

$$
\begin{cases}
-\frac{\partial F}{\partial p_k} - \lambda_k + \mu = 0, \forall k, \\
\lambda_k = 0 \text{ or } p_k = 0, \forall k, \\
\mu = 0 \text{ or } \sum_{j=1}^{K} p_j = P_{\max}.
\end{cases}
\tag{21}
$$

where λ_k is the Lagrange multiplier associated with $p_k \geq 0$, and μ is the Lagrange multiplier associated with $\sum_{j=1}^{K} p_j \leq P_{\max}$.

At the optimal solution, if at least one entry of $\mathbf{p}$ is greater than zero,[5] we have $i_1 < i_2 < \ldots < i_M \in \{1, 2, \ldots, K\}$, $M \geq 1$, $p_{i_1} > 0, p_{i_2} > 0, \ldots, p_{i_M} > 0$. $\{i_1, i_2, \ldots, i_M\}$ is an arbitrary non-empty subset of $\{1, 2, \ldots, K\}$. Then, following the KKT condition, we have $\lambda_{i_1} = \lambda_{i_2} = \ldots = \lambda_{i_M} = 0$ and thus

$$
\frac{\partial F}{\partial p_{i_1}} = \frac{\partial F}{\partial p_{i_2}} = \ldots = \frac{\partial F}{\partial p_{i_M}}.
\tag{22}
$$

If $\mu \neq 0$, we have

$$
p_{i_1} + p_{i_2} + \ldots + p_{i_{M-1}} = P_{\max}.
\tag{23}
$$

Otherwise, $\mu = 0$, and we have

$$
\frac{\partial F}{\partial p_{i_1}} = \frac{\partial F}{\partial p_{i_2}} = \ldots = \frac{\partial F}{\partial p_{i_M}} = 0.
\tag{24}
$$

As a consequence, either (23) or (24) stands.

Substituting $p_j = 0$, $j \neq i_1, \ldots, i_M$, into (20), and ignoring those lines regarding to $\frac{\partial F}{\partial p_j}$, $j \neq i_1, \ldots, i_M$, we have

$$
\begin{cases}
\frac{\partial F}{\partial p_{i_1}} = \frac{g_{i_1} Q_{i_1}}{p_{i_1} g_{i_1} + \eta} - \frac{g_{i_2} Q_{i_2}}{p_{i_1} g_{i_2} + \eta} \\
\quad + \frac{g_{i_2} Q_{i_2}}{(p_{i_1} + p_{i_2}) g_{i_2} + \eta} - \frac{g_{i_3} Q_{i_3}}{(p_{i_1} + p_{i_2}) g_{i_3} + \eta} + \ldots \\
\quad + \frac{g_{i_{M-1}} Q_{i_{M-1}}}{(p_{i_1} + \ldots + p_{i_{M-1}}) g_{i_{M-1}} + \eta} - \frac{g_{i_M} Q_{i_M}}{(p_{i_1} + \ldots + p_{i_{M-1}}) g_{i_M} + \eta} \\
\quad + \frac{g_{i_M} Q_{i_M}}{(p_{i_1} + \ldots + p_{i_M}) g_{i_M} + \eta} - Z, \\
\frac{\partial F}{\partial p_{i_2}} = \frac{g_{i_2} Q_{i_2}}{(p_{i_1} + p_{i_2}) g_{i_2} + \eta} - \frac{g_{i_3} Q_{i_3}}{(p_{i_1} + p_{i_2}) g_{i_3} + \eta} + \ldots \\
\quad + \frac{g_{i_{M-1}} Q_{i_{M-1}}}{(p_{i_1} + \ldots + p_{i_{M-1}}) g_{i_{M-1}} + \eta} - \frac{g_{i_M} Q_{i_M}}{(p_{i_1} + \ldots + p_{i_{M-1}}) g_{i_M} + \eta} \\
\quad + \frac{g_{i_M} Q_{i_M}}{(p_{i_1} + \ldots + p_{i_M}) g_{i_M} + \eta} - Z, \\
\ldots \\
\frac{\partial F}{\partial p_{i_{M-1}}} = \frac{g_{i_{M-1}} Q_{i_{M-1}}}{(p_{i_1} + \ldots + p_{i_{M-1}}) g_{i_{M-1}} + \eta} - \frac{g_{i_M} Q_{i_M}}{(p_{i_1} + \ldots + p_{i_{M-1}}) g_{i_M} + \eta} \\
\quad + \frac{g_{i_M} Q_{i_M}}{(p_{i_1} + \ldots + p_{i_M}) g_{i_M} + \eta} - Z, \\
\frac{\partial F}{\partial p_{i_M}} = \frac{g_{i_M} Q_{i_M}}{(p_{i_1} + \ldots + p_{i_M}) g_{i_M} + \eta} - Z.
\end{cases}
\tag{25}
$$

By combining (22) into (25), we have

$$
\begin{cases}
\frac{g_{i_1} Q_{i_1}}{p_{i_1} g_{i_1} + \eta} - \frac{g_{i_2} Q_{i_2}}{p_{i_1} g_{i_2} + \eta} = 0, \\
\frac{g_{i_2} Q_{i_2}}{(p_{i_1} + p_{i_2}) g_{i_2} + \eta} - \frac{g_{i_3} Q_{i_3}}{(p_{i_1} + p_{i_2}) g_{i_3} + \eta} = 0, \\
\ldots \\
\frac{g_{i_{M-1}} Q_{i_{M-1}}}{(p_{i_1} + \ldots + p_{i_{M-1}}) g_{i_{M-1}} + \eta} - \frac{g_{i_M} Q_{i_M}}{(p_{i_1} + \ldots + p_{i_{M-1}}) g_{i_M} + \eta} = 0,
\end{cases}
\tag{26}
$$

[4]Please note that the KKT condition does not guarantee global optimality, but global optimality implies the KKT condition.

[5]If all entries of the optimal $\mathbf{p}$ equal to zero, $\forall k \leq K$, the values of $p_1 + p_2 + \ldots + p_k$ are always 0, so that the Lemma 1 is trivially true. This scenario can be ignored in this proof.

which leads to

$$
\begin{cases}
p_{i_1} = \frac{\eta(g_{i_2}Q_{i_2}-g_{i_1}Q_{i_1})}{g_{i_1}g_{i_2}(Q_{i_1}-Q_{i_2})}, \\
p_{i_1} + p_{i_2} = \frac{\eta(g_{i_3}Q_{i_3}-g_{i_2}Q_{i_2})}{g_{i_2}g_{i_3}(Q_{i_2}-Q_{i_3})}, \\
\cdots \\
p_{i_1} + p_{i_2} + \ldots + p_{i_{M-1}} = \frac{\eta(g_{i_M}Q_{i_M}-g_{i_{M-1}}Q_{i_{M-1}})}{g_{i_{M-1}}g_{i_M}(Q_{i_{M-1}}-Q_{i_M})}.
\end{cases}
\tag{27}
$$

By combining (23) and (24) into the last line of (25), we have either

$$
\frac{\partial F}{\partial p_{i_M}} = \frac{g_{i_M}Q_{i_M}}{(p_{i_1}+\ldots+p_{i_M})g_{i_M}+\eta} - Z = 0,
\tag{28}
$$

or

$$
p_{i_1} + \ldots + p_{i_M} = P_{\max}.
\tag{29}
$$

Therefore, by combining (27), (28), and (29), $p_{i_1} + p_{i_2} + \ldots + p_{i_m}$, $m \le M$ can only possibly be one from the following values in $\mathcal{E}(i_1, i_2, \ldots, i_M) = \left\{ \frac{\eta(g_{i_2}Q_{i_2}-g_{i_1}Q_{i_1})}{g_{i_1}g_{i_2}(Q_{i_1}-Q_{i_2})}, \frac{\eta(g_{i_3}Q_{i_3}-g_{i_2}Q_{i_2})}{g_{i_2}g_{i_3}(Q_{i_2}-Q_{i_3})}, \ldots, \frac{\eta(g_{i_M}Q_{i_M}-g_{i_{M-1}}Q_{i_{M-1}})}{g_{i_{M-1}}g_{i_M}(Q_{i_{M-1}}-Q_{i_M})}, \frac{Q_{i_M}}{Z} - \frac{\eta}{g_{i_M}}, P_{\max} \right\}$.

Finally, $\{i_1, i_2, \ldots, i_M\}$ could be an arbitrary subset of $\{1, 2, \ldots, K\}$. Thus, $p_1 + p_2 + \ldots + p_k$ can only possibly be 0 or one value in $\bigcup_{\text{all possible } i_1,i_2,\ldots,i_M} \mathcal{E}(i_1, i_2, \ldots, i_M)$, which is equivalent to $\left\{ 0; \frac{\eta(g_jQ_j-g_iQ_i)}{g_ig_j(Q_i-Q_j)}, \forall i, j, 1 \le i < j \le K; \frac{Q_i}{Z} - \frac{\eta}{g_i}, \forall i, 1 \le i \le K; P_{\max} \right\}$. This completes the proof. $\square$

Lemma 1 shows that when the optimality of Problem PA is achieved, the value of $p_1 + p_2 + \ldots + p_k, \forall k \le K$, must be one from the candidate values in $\mathcal{E}$. This is because if a value of $p_1 + p_2 + \ldots + p_k, \forall k \le K$, is not in $\mathcal{E}$, the KKT condition is not satisfied and **p** is not an optimal solution. Such property greatly decreases the searching space for the optimal solution, which will then be employed as a key in the DPPA algorithm.

Please note that if a value in $\mathcal{E}$ is greater than $P_{\max}$ or smaller than 0, $p_1 + p_2 + \ldots + p_k, \forall k \le K$, cannot be equal to that value. Therefore, we can eliminate such values from $\mathcal{E}$. Let $\widetilde{\mathcal{E}} \triangleq \{e \in \mathcal{E} | 0 \le e \le P_{\max}\}$. We have the following corollary

Corollary 1: If the optimality of Problem PA is achieved, the value of $p_1 + p_2 + \ldots + p_k, \forall k \le K$, can only be one from those in $\widetilde{\mathcal{E}}$.

In the rest of this paper, let $L = |\widetilde{\mathcal{E}}|$ represent the number of elements in $\widetilde{\mathcal{E}}$. $L \le \frac{K(K-1)}{2} + K + 2$ since there are $\frac{K(K-1)}{2} + K + 2$ elements in $\mathcal{E}$ by definition.

C. Incremental Structure

In this subsection, we present the incremental structure in the optimal solution, which will help to construct the Bellman equation to be used in the dynamic programming.

For convenience, we define

$$
f_1(p_1) \triangleq Q_1 \log\left(1 + \frac{p_1 g_1}{\eta}\right) - Z \cdot p_1,
\tag{30}
$$

$$
f_2(p_1, p_2) \triangleq Q_2 \log\left(1 + \frac{p_2 g_2}{p_1 g_2 + \eta}\right) - Z \cdot p_2,
\tag{31}
$$

$$
\cdots
$$

$$
f_K(p_1 + p_2 + \ldots + p_{K-1}, p_K)
$$
$$
\triangleq Q_K \log\left(1 + \frac{p_K g_K}{\sum_{j=1}^{K-1} p_j g_K + \eta}\right) - Z \cdot p_K.
\tag{32}
$$

Let

$$
G_k(p_1, p_2, \ldots, p_k)
$$
$$
\triangleq f_1(p_1) + f_2(p_1, p_2) + \ldots
$$
$$
+ f_k(p_1 + p_2 + \ldots + p_{k-1}, p_k), \forall k.
\tag{33}
$$

Then, we have

$$
G_k(p_1, p_2, \ldots, p_k) = G_{k-1}(p_1, p_2, \ldots, p_{k-1})
$$
$$
+ f_k(p_1 + p_2 + \ldots + p_{k-1}, p_k).
\tag{34}
$$

Please note that $G_K(p_1, p_2, \ldots, p_K) = F(p_1, p_2, \ldots, p_K)$, the objective function of Problem PA.

We notice that if we want to obtain the optimal solution to PA, we can solve them by gradually improving the objective function, using the following incremental structure.

Lemma 2 (Incremental Structure): If $p_1^*, p_2^*, \ldots, p_k^*$, $k \ge 2$, is an optimal solution to the following problem

$$
\max_{p_1, p_2, \ldots, p_k} G_k(p_1, p_2, \ldots, p_k),
\tag{35a}
$$
$$
\text{subject to} \quad p_1 + p_2 + p_3 + \ldots + p_k = A,
\tag{35b}
$$
$$
p_j \ge 0, \quad \forall j = 1, 2, \ldots k,
\tag{35c}
$$

where A is some non-negative value, then, $p_1^*, p_2^*, \ldots, p_{k-1}^*$ is an optimal solution to the following problem

$$
\max_{p_1, p_2, \ldots, p_{k-1}} G_{k-1}(p_1, p_2, \ldots, p_{k-1}),
\tag{36a}
$$
$$
\text{subject to} \quad p_1 + p_2 + p_3 + \ldots + p_{k-1}
\tag{36b}
$$
$$
= p_1^* + p_2^* + p_3^* + \ldots + p_{k-1}^*,
$$
$$
p_j \ge 0, \quad \forall j = 1, 2, \ldots k - 1.
\tag{36c}
$$

Proof: Suppose $p_1^*, p_2^*, \ldots, p_{k-1}^*$ is not an optimal solution to problem (36a)–(36c), then we can find another solution $p_1^{**}, p_2^{**}, \ldots, p_{k-1}^{**}$ leading to a even larger value of the object function (36a), $f_1(p_1^{**}) + f_2(p_1^{**}, p_2^{**}) + \ldots + f_k(p_1^{**} + p_2^{**} + \ldots + p_{k-2}^{**}, p_{k-1}^{**}) > f_1(p_1^*) + f_2(p_1^*, p_2^*) + \ldots + f_k(p_1^* + p_2^* + \ldots + p_{k-2}^*, p_{k-1}^*)$.

Please note that $p_1^{**}, p_2^{**}, \ldots, p_{k-1}^{**}, p_k^*$ is also a feasible solution to (35a)–(35c), however, such solution leads to a greater value in the objective function in (35a), i.e., $f_1(p_1^{**}) + f_2(p_1^{**}, p_2^{**}) + f_2(p_1^{**} + p_2^{**}, p_3^{**}) + \ldots + f_k(p_1^{**} + p_2^{**} + \ldots + p_{k-1}^{**}, p_k^*) > f_1(p_1^*) + f_2(p_1^*, p_2^*) + f_2(p_1^* + p_2^*, p_3^*) + \ldots + f_k(p_1^* + p_2^* + \ldots + p_{k-1}^*, p_k^*)$, leading to a contradiction. Through this contradiction, Lemma 2 is proved. $\square$

Lemma 2 implies that in order to solve the optimization problem with the objective function $G_k(\cdot)$, we can first solve a reduced subproblem with the objective function $G_{k-1}(\cdot)$. Therefore, the original problem with objective function $G_K(\cdot)$ can then be solved recursively, through solving optimization problems with objective functions

$G_{K-1}(\cdot), G_{K-2}(\cdot), \ldots, G_1(\cdot)$. In what follows, we realize such idea by detailed derivations: We first derive the Bellman equation showing the recursion of the optimality. Then, a dynamic programming approach is proposed to solve the Problem PA.

D. Recursion of Optimality: Bellman Equation

In the next step, we aim to solve the Problem PA via a dynamic programming approach. Intrinsically, dynamic programming is a mathematical optimization method. The original problem is broken down into simpler sub-problems in a recursive manner, so that the original problem can be solved optimally through recursively finding the optimal solutions to the sub-problems. The mathematical relation between the optimality of the problem and its sub-problems is characterized by the Bellman equation [38, Ch. 15].

As a prerequisite for dynamic programming, the Bellman equation regarding to the optimal solution to Problem PA is derived in this subsection, which is achieved through taking the advantage of the structures derived by Lemmas 1 and 2.

We define the Subproblem-l-k (SP-l-k) as follows

$$\max_{p_1,p_2,\ldots,p_k} \quad G_k(p_1, p_2, \ldots, p_k), \tag{37a}$$

$$\text{subject to} \quad p_1 + p_2 + p_3 + \ldots + p_k = \pi_l, \tag{37b}$$

$$p_j \geq 0, \quad \forall j = 1, 2, \ldots k, \tag{37c}$$

$$p_1 + p_2 + \ldots + p_{k'} \in \widetilde{\mathcal{E}}, \quad \forall k' \leq k, \tag{37d}$$

where π_l is the lth smallest element in $\widetilde{\mathcal{E}}$. $\{\pi_1, \pi_2, \ldots, \pi_L\}$ is the sorted version (in ascending order) of $\widetilde{\mathcal{E}}$. We define $H(l, k)$ as the optimal G_k value of SP-l-k, and let $\mathbf{p}_{opt}(l, k) = (p_{opt,1}(l, k), \ldots, p_{opt,k}(l, k))$ denote the optimal solution. Following Corollary 1, since $p_1 + p_2 + p_3 + \ldots + p_K$ is one from the values in $\{\pi_1, \pi_2, \ldots, \pi_L\}$, the optimal solution to Problem PA is the best one among the optimal solutions to SP-1-K, SP-2-K, ..., SP-L-K.

Then, we characterize $H(l, k)$ through analyzing $H(l', k-1)$. When SP-l-k is optimized, $p_1 + p_2 + p_3 + \ldots + p_{k-1}$ must be equal to some $\pi_{l'} \in \widetilde{\mathcal{E}}$, $l' \leq l$. Given l', due to Lemma 2, the following SP-l'-$(k-1)$ must be optimized first

$$\max_{p_1,p_2,\ldots,p_{k-1}} \quad G_{k-1}(p_1, p_2, \ldots, p_{k-1}), \tag{38a}$$

$$\text{subject to} \quad p_1 + p_2 + p_3 + \ldots + p_{k-1} = \pi_{l'}, \tag{38b}$$

$$p_j \geq 0, \quad \forall j = 1, 2, \ldots k-1, \tag{38c}$$

$$p_1 + p_2 + \ldots + p_{k'} \in \widetilde{\mathcal{E}}, \quad \forall k' \leq k-1. \tag{38d}$$

However, at this stage, we do not know the value l', so that we need to examine all possible l' values. By solving all possible SP-l'-$(k-1)$, $\forall l' \leq l$, we can find the best one among them to achieve the optimal SP-l-k.

Given that SP-l'-$(k-1)$ is optimized (it implies $p_1 + p_2 + p_3 + \ldots + p_{k-1} = \pi_{l'}$), and given the constraint $p_1 + p_2 + p_3 + \ldots + p_k = \pi_l$ in SP-l-k, p_k is fixed at $\pi_l - \pi_{l'}$. Then the value of the objective function of SP-l-k becomes $H(l', k-1) + f_k(\pi_{l'}, \pi_l - \pi_{l'})$ in this scenario. By examining all possible l', $\forall l' \leq l$, $H(l, k)$ is the largest one among all possible $H(l', k-1) + f_k(\pi_{l'}, \pi_l - \pi_{l'})$ values, $l' \leq l$.

As a consequence, we have the following Bellman equation showing the relationship between the optimal solution to SP-l-k and its subproblems for $k \geq 2$

$$H(l, k) = \max_{l'=1,2,\ldots,l} (H(l', k-1) + f_k(\pi_{l'}, \pi_l - \pi_{l'})). \tag{39}$$

When $k = 1$, the optimal value of the objective of SP-l-1 can be simply derived as $H(l, 1) = f_1(\pi_l)$ following definition.

The optimal solution to Problem PA is $\max[H(1, K), H(2, K), \ldots, H(L, K)]$, which can be derived through recursively finding the optimal solutions to $H(l, K-1), H(l, K-2), \ldots, H(l, 1), \forall l$, by employing dynamic programming. In the next subsection, we will formally propose the Dynamic Programming based Power Allocation (DPPA) algorithm to solve Problem PA.

E. Dynamic Programming Based Power Allocation (DPPA)

In this subsection, we propose the DPPA formally described in Algorithm 1 through the derived the Bellman equation given in (39).

In Lines 1–16 of Algorithm 1, we construct the set $\widetilde{\mathcal{E}}$. In Lines 17–26, we derive the optimal $H(k, l)$ values via dynamic programming approach. Finally, the optimal solution to Problem PA is derived in Lines 27–29, since the optimal solution to Problem PA is the best one among the optimal solutions to SP-1-K, SP-2-K, ..., SP-L-K.

F. Complexity Analysis

The computational complexity of the above DPPA is characterized as follows. In Lines 20–26, there are two levels of "for" loops. The outer level (Line 20) introduces a computational complexity of $O(K)$, and the inner level (Line 21) introduces a computational complexity of $O(L)$. Within the inner "for" loop (Lines 22–23), we search for the maximum term among l terms, and $l \leq L$, introducing another level of computational complexity of $O(L)$. Therefore, the overall computational complexity is $O(KL^2)$. L is in the scale of $O(K^2)$ according to Corollary 1, so that the overall computational complexity is $O(K^5)$, which is polynomial.

Note that for a general non-convex optimization problem, there are no standard methods to find a globally optimal solution within a polynomial computational complexity. In order to derive a globally optimal solution to Problem PA, one alternative method is to study the KKT condition of the optimization problem and exhaustively search for all the solutions satisfying the KKT condition. Then, the best solution among them is selected as the globally optimal solution.

By this method, the KKT condition is shown in (21), and then we need to find all solutions satisfying (21). According to the second line of (21), for each user k, we need to consider two possibilities: (1) $\lambda_k = 0$ and (2) $p_k = 0$ for this user. Since there are K users, there are 2^K possibilities. In the third line of (21), we need to consider another set of two possibilities. In sum, we need to exhaustively search for 2×2^K possibilities to find all solutions satisfying (21), in order to find the best one among them. Therefore, the computational complexity is $O(2^K)$, non-polynomial, without our proposed DPPA algorithm.

Algorithm 1 Dynamic Programming based Power Allocation

---

1 Find all possible $p_1 + p_2 + \ldots + p_k, k \leq K$, values and put them in $\widetilde{\mathcal{E}}$;

2 $\widetilde{\mathcal{E}} \leftarrow \{0, P_{\max}\}$;

3 **for** $i \leftarrow 1$ **to** K **do**

4   **for** $j \leftarrow i+1$ **to** K **do**

5         $p_a \leftarrow \frac{\eta \ (g_j Q_j - g_i Q_i)}{g_i g_j (Q_i - Q_j)}$;

6   **if** $0 < p_a < P_{\max}$ **then**

7   Put p_a in $\mathcal{E}$;

8   **end**

9   **end**

10     $p_a \leftarrow \frac{Q_i}{Z} - \frac{\eta}{g_i}$;

11   **if** $0 < p_a < P_{\max}$ **then**

12   Put p_a in $\widetilde{\mathcal{E}}$;

13   **end**

14 **end**

15 $L \leftarrow |\widetilde{\mathcal{E}}|$, i.e., the size of $\widetilde{\mathcal{E}}$;

16 Sort all elements in $\widetilde{\mathcal{E}}$ in ascending order $\pi_1, \pi_2, \ldots, \pi_L$;

17 Dynamic programming using Bellman equation;

18 Set $H(l,k) \leftarrow 0, \forall k = 1, 2, \ldots, K, \forall l = 1, 2, \ldots, L$;

19 Set $H(l,1) \leftarrow f_1(\pi_l)$, $\mathbf{p}_{opt}(l,1) \leftarrow (\pi_l)$, $\forall l = 1, 2, \ldots, L$;

20 **for** $k \leftarrow 2$ **to** K **do**

21   **for** $l \leftarrow 1$ **to** L **do**

22         $H(l,k) \leftarrow$ $\max_{l'=1,2,\ldots,l} \left[H(l',k-1) + f_k(\pi'_l, \pi_l - \pi'_l) \right]$;

23         $l_0 \leftarrow$ $\arg\max_{l'=1,2,\ldots,l} \left[H(l',k-1) + f_k(\pi'_l, \pi_l - \pi'_l) \right]$;

24         $\mathbf{p}_{opt}(l,k) \leftarrow (\mathbf{p}_{opt}(l_0, k-1), \pi_l - \pi_{l_0})$; append $\pi_l - \pi_{l_0}$ in the end;

25   **end**

26 **end**

27 $H_{opt} \leftarrow \max_{l=1,2,\ldots,L} H(l,K)$; Optimal value of objective function of Problem PA;

28 $l_0 \leftarrow \arg\max_{l=1,2,\ldots,L} H(l,K)$;

29 $\mathbf{p}_{opt} \leftarrow \mathbf{p}_{opt}(l_0, K)$; Optimal power allocation of Problem PA.

IV. SIMULATION PERFORMANCE EVALUATION

In this section, we present simulation based performance evaluation of the proposed long-term resource allocation framework for NOMA. We first introduce the simulation setup and four benchmark schemes, and then demonstrate the performance gain of our DPPA-based NOMA scheme over these benchmark schemes. For simplification, the NOMA scheme proposed in this paper is referred to as NOMA-OPT in the rest of this section.

A. Simulation Setup

The channel gain from the BS to each user is computed as $\frac{1}{D^\alpha} \cdot g_0$, where D^α is the pathloss, D (in meter) is the distance between the BS to the user, $\alpha = 4$ is the pathloss exponent, and g_0 is independently exponentially distributed

with a unit mean (corresponding to a normalized Rayleigh fading). We set $W = 20$ MHz and $\tau = 50$ ms. The utility function of each user is set to be ln(data rate in Mbit per timeslot). Each simulation point is averaged over 50000 timeslots. We set $P_{\max} = 33$ dBm, $P_{\text{mean}} = 30$ dBm, $R_{\max} = 15$ Mbits, and the noise $\eta = -87$ dBm. The initial queue backlogs for all users are 0.

B. Benchmark Schemes

In this section, we compare the proposed NOMA-OPT scheme with four benchmark schemes as follows.

1) Time-Sharing Orthogonal Multiple Access (OMA) Scheme: In this scheme, different users are operated on equally partitioned time durations in each timeslot. For fairness, the OMA scheme has the same queueing model as the NOMA-OPT scheme: Problem PO is still the initial objective to be solved. The same Lyapunov optimization analysis can be performed, and PO is converted to a single-timeslot optimization, which can then be decomposed into Problems RC-i and Problem PA. The only difference is that the objective function of Problem PA is replaced by $\max_{\mathbf{p}} \frac{1}{K} Q_1 \log \left(1 + \frac{p_1 g_1}{\eta} \right) + \frac{1}{K} Q_2 \log \left(1 + \frac{p_2 g_2}{\eta} \right) + \ldots + \frac{1}{K} Q_K \log \left(1 + \frac{p_K g_K}{\eta} \right) - Z \left(\sum_{i=1}^{K} p_i \right)$. We can verify that Problem PA for OMA is a convex optimization problem, which can be solved by a standard method (e.g., interior point method). The derived optimal solution is adopted in each timeslot.

2) Single User Serving (Single) Scheme: In this scheme, only one single user is served in each timeslot. For fairness, the Single scheme has the same queueing model as the NOMA-OPT scheme. The objective function of Problem PA does not change, but an additional constraint that at most one of $p_1, p_2, \ldots, p_K$ can be positive is added. Through this way, the resultant power is allocated to one "best" user in each timeslot.[6]

3) NOMA With Equal Power Allocation (NOMA-EQ) Scheme: In this scheme, NOMA is employed, but power allocation is not optimized. In each timeslot, power is equally allocated to all users, and the total power allocated is equal to P_{mean}. For fairness, the NOMA-EQ scheme has the same queueing model as the NOMA-OPT scheme.

4) NOMA With Power Allocation Proportional to Queue Backlog (NOMA-Pro-Q) Scheme: In this scheme, NOMA is employed, but power allocation is not optimized. In each timeslot, the power allocated to each user is proportional to the queue backlog of that user ($Q_i(t)$), and the total power allocated is equal to P_{mean}. The NOMA-EQ scheme has the same queueing model as the NOMA-OPT scheme.

---

[6]Note that "Single" is not optimal to Problem PA. If $g_i > g_j$, then serving user i instead of user j will lead to a better solution. If $Q_i > Q_j$, we should give a higher priority to serve user i since the queue backlog at user i is larger. If $g_i > g_j$ and $Q_i > Q_j$, serving user i but not j is optimal. However, if we have $g_i > g_j$ and $Q_i < Q_j$, it is not straightforward to decide which user to serve, and how much power should be allocated to users i and j. We can only solve the problem optimally through our proposed NOMA-OPT solution, which is one of our core contributions of this work.
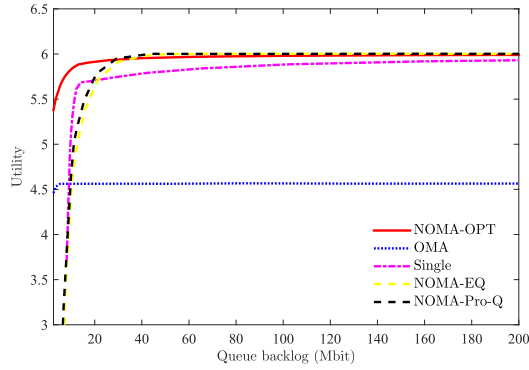
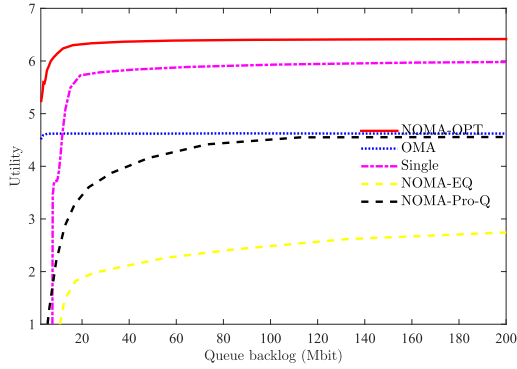Fig. 2.　Queue backlog versus network utility in Scenario 1.

Fig. 3.　Queue backlog versus network utility in Scenario 2.

C. Tradeoff Between Utility and Queue Backlog

First, we study the tradeoff between the long-term user utility and the average queue backlog $Q_i(t)$, when we adjust the value of V from 10^{-1} to 10^3. Recall that V is a tunable value that achieves an $[O(V), O(\frac{1}{V})]$ tradeoff between the queue backlogs and utility, shown in (15)–(16). We study two scenarios. There are $K = 5$ users in both scenarios. In Scenario 1, all of the five users are 100 meters from the BS; and in Scenario 2, the five users are 60, 80, 100, 120, and 140 meters from the BS.

The tradeoff performance of NOMA-OPT and four benchmark schemes in Scenarios 1 and 2 is shown in Figs. 2 and 3. First, we observe that when the queue backlog increases, the user utility increases fast at the beginning, but then gradually saturates. This matches our analysis that there is an $[O(V), O(\frac{1}{V})]$ tradeoff. Please note that here $O(V)$ indicates the size of queue backlog, and $O(\frac{1}{V})$ indicates the gap between the utility under V and the maximum possible utility. Therefore, the utility asymptotically approaches the maximum possible utility when we keep increasing the queue backlog.

Second, in both figures, we observe that the performance of NOMA-OPT is much better than those of OMA and Single. Given an arbitrary queue backlog, the overall utility of NOMA-OPT is higher than those of Single and OMA. This matches our expectation: NOMA-OPT outperforms OMA because NOMA can better utilize the radio resource to achieve better data rates for all users. NOMA-OPT outperforms Single because Single gives a sub-optimal solution to Problem PA in each timeslot (since an additional constraint that at most one user can be served is added in Single).
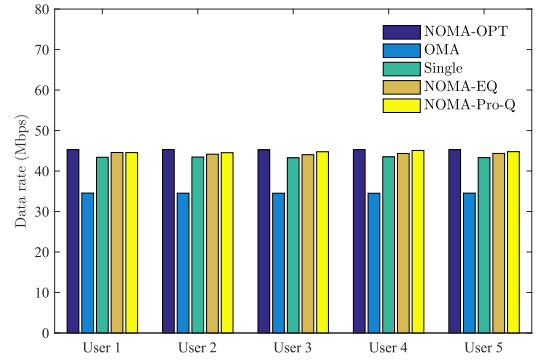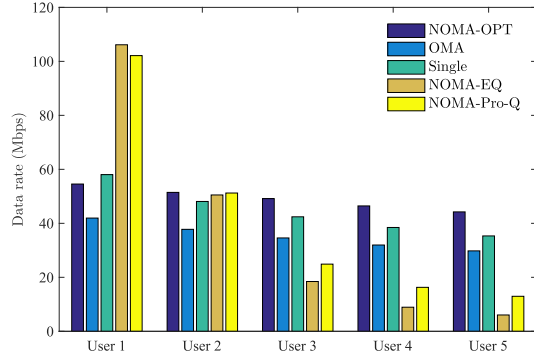

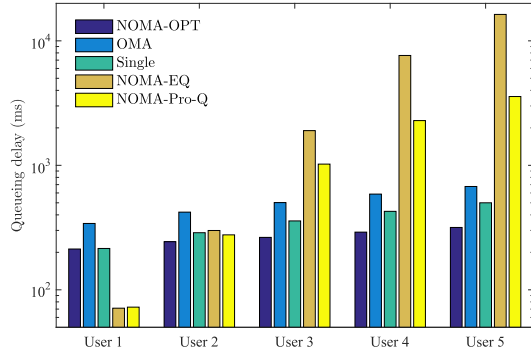
(a) Data rate.
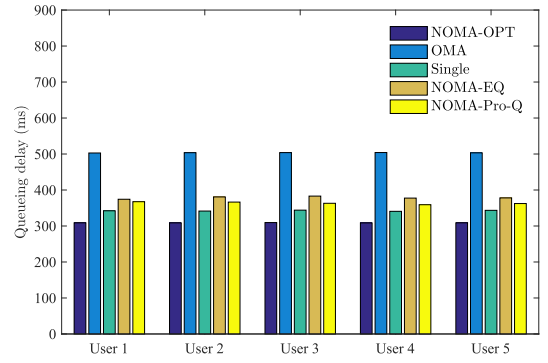
(b) Queueing delay.
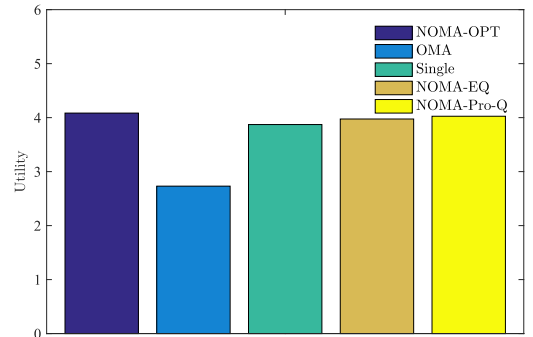
(c) Utility.

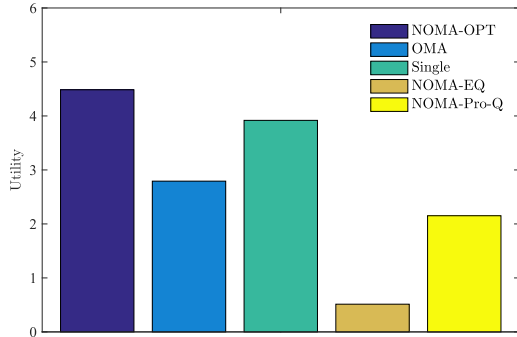Fig. 4.　Performance comparison in Scenario 1.

Third, we observe that in Fig. 2, the performances of NOMA-OPT, NOMA-EQ, and NOMA-Pro-Q schemes are close to each other. This is because in Scenario 1, the distances from the five users to the BS are the same, so that the five users are symmetric. Equal power allocation is already a sufficiently good solution to these five symmetric users. In this scenario, for NOMA-Pro-Q scheme, the five users experience almost the same queue backlogs, so that NOMA-Pro-Q is almost equivalent to NOMA-EQ. However, in Fig. 3 when the distances from the five users to the BS are different (which is most likely in reality), NOMA-OPT substantially outperforms NOMA-EQ and NOMA-Pro-Q. This is because NOMA-EQ does not allocate sufficient power to users far from the BS, so that these users experience very low data rate, leading to low log utility. Compared with NOMA-EQ, NOMA-Pro-Q improves the log utility by allocating more power to users

(a) Data rate.

(b) Queueing delay.

(c) Utility.
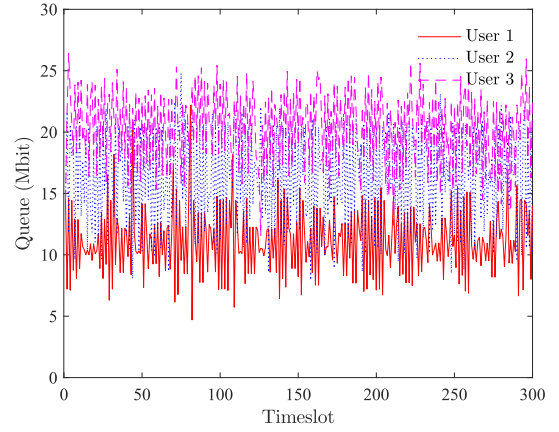
Fig. 5. Performance comparison in Scenario 2.

Fig. 6. Evolution of queue backlogs in Scenario 3.

the five schemes. $V = 30$ in this subsection. In Figs. 4 and 5, we show the data rate, average queueing delay, and the overall utility in subfigures (a), (b), and (c), respectively. In Scenario 1, NOMA-OPT slightly improves the data rates and queueing delays for all users compared with Single, NOMA-EQ, and NOMA-Pro-Q, but substantially improves data rates and queueing delays for all users compared with OMA. In Scenario 2, the advantage of NOMA-OPT is much more substantial. In particular, we observe

1) Compared with OMA, the data rates and queueing delays for all users are substantially improved.
2) Compared with Single, the data rate of user 1 is slightly smaller. This gives way to the performance gains of users 2, 3, 4, and 5, and thus the overall log utility shown in Fig. 5(c). In addition, the delay performance of users 4 and 5 are substantially improved (Note that the y-axis of Fig. 5(b) is shown in log scale). This is because the performance gain of NOMA-OPT stems from the optimality in solving Problem PA (while Single gives a non-optimal solution).
3) NOMA-EQ and NOMA-Pro-Q only provide large data rate and small delay to the closest user, while the data rate and delay performance of far users are very poor. Therefore, without our proposed NOMA-OPT approach, the performance NOMA could be unsatisfactory.
4) NOMA-OPT brings the highest performance enhancement to user 5 (the farthest user) compared with all the other four schemes. It suggests that edge users will get more benefits if NOMA-OPT is adopted.

E. Evolution of Queue Backlogs

In this subsection, we investigate the evolution of the queue backlogs and understand how they influence the system fairness, as shown in Fig. 6. For illustration purpose, we study the following Scenario 3: There are $K = 3$ users, and they are 20, 100, and 200 meters from the BS. Obviously, their distances to the BS and thus channel gains are quite different in this scenario. $V = 50$ in this subsection.

Through the NOMA-OPT approach, the data rates of the three users are 215.65, 137.94, and 105.82 Mbps respectively. Even if the channel gain of the furthest user is 40 dB smaller than that of the closest user (on average), it can still achieve

farther away. However, the improved performance is still much worse than that of NOMA-OPT.

Finally, we notice that without wisely devising the power allocation scheme for NOMA, the benefit of NOMA cannot be realized. The performances of NOMA-EQ and NOMA-Pro-Q are even much worse than that of OMA in Fig. 3. Such observation further demonstrates the significance of our proposed NOMA-OPT scheme.

D. User Data Rate and Queueing Delay

We also study the average user data rate and average queueing delay[7] of the five users in Scenarios 1 and 2, under

---

[7]It is the average time of a data bit staying in queue $Q_i(t)$. In the simulation, we track the timestamps at which each data bit enters and leaves the queue $Q_i(t)$, and compute the time difference. The result is averaged over all data bits in the simulation.
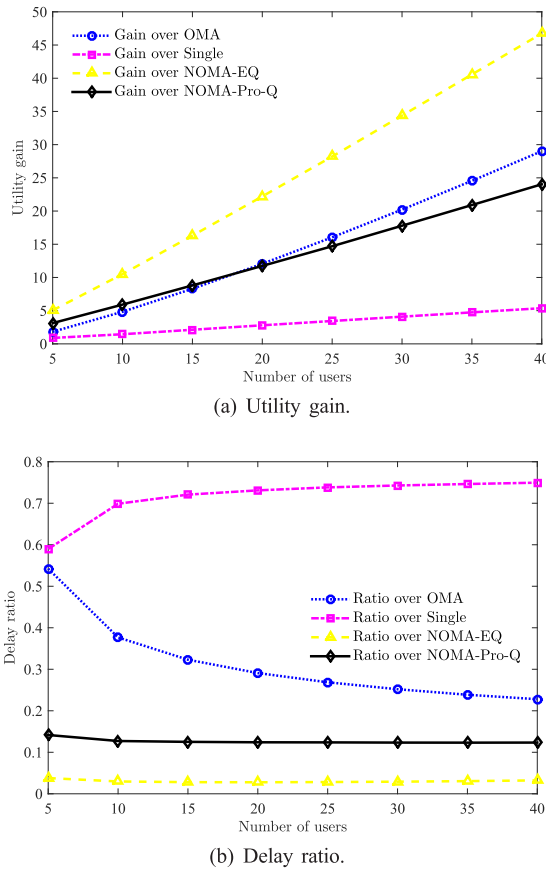
Fig. 7. Performance comparison under different numbers of users.

48.8% of data rate of the closest user, suggesting that the system fairness is well protected. This is aligned with the fact that the queue backlogs of the three users fluctuate around 11, 16, and 20 respectively: user 3 is usually associated with a larger queue backlog and thus has a higher weight (priority) in the optimization problem in (19a).

F. Performance Gain Under Different Numbers of Users

Finally, we investigate the performance gain of NOMA-OPT over the four benchmark schemes under variable numbers of users. $V = 20$ in this subsection. We consider scenarios where the number of users increases from 5 to 40. In each case, users are lined from 50 meters to 150 meters from the BS with equal intervals. We study the utility gain of NOMA-OPT compared with each of the benchmark schemes (i.e., utility of NOMA-OPT minus utility of the benchmark scheme) and the delay ratio of NOMA-OPT over each of the benchmark schemes (i.e., average queueing delay of NOMA-OPT over average queueing delay of the benchmark scheme). The results shown in Fig. 7 suggest that NOMA-OPT brings substantial performance gain, in terms of both data rate and delay, under a variety of user numbers in the system. In addition, the utility gain increases as the number of user increases. This is because the original Problem PO is solved asymptotically optimally by our proposed NOMA-OPT, but all of the benchmark schemes give non-optimal solutions. The more users in the system, a non-optimal solution has a higher chance to make a "poor"

decision, and thus causes larger performance loss compared with NOMA-OPT.

V. CONCLUSIONS

In this paper, we developed a long-term resource allocation framework for a downlink multi-user non-orthogonal multiple access (NOMA) network, which jointly optimizes the rate control at the network layer as well as the power allocation at the physical layer to maximize the long-term network utility, subject to several practical long-term and short-term constraints on the power consumption and user queue stability. By resorting to the Lyapunov optimization theory, we attained the asymptotically optimal solution to the formulated long-term network utility maximization problem. To achieve this, we converted the utility maximization problem into a series of short-term rate control and power allocation problems to be optimized in each timeslot and successfully resolved these short-term problems by leveraging the special structures of the objective functions. Simulation results were presented to compare the performance of the proposed NOMA method with those of four benchmark schemes, including OMA and non-optimal NOMA schemes, under the same setups. The simulation results showed that compared to these benchmark schemes, the proposed NOMA method can achieve a higher network throughput, lower data delay, and better user fairness.

REFERENCES

[1] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[2] G. Wunder *et al.*, "5GNOW: Non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 97–105, Feb. 2014.
[3] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi, "Superposition coding strategies: Design and experimental evaluation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2628–2639, Jul. 2012.
[4] 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.
[5] Z. Ding *et al.* (2015). "Application of non-orthogonal multiple access in LTE and 5G networks." [Online]. Available: https://arxiv.org/pdf/1511.08610.pdf
[6] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721–742, 2nd Quart., 2016.
[7] Z. Wei, J. Yuan, D. W. K. Ng, M. Elkashlan, and Z. Ding. (2016). "A survey of downlink non-orthogonal multiple access for 5G wireless communication networks." [Online]. Available: https://arxiv.org/abs/1609.01856
[8] H. Lee, S. Kim, and J. H. Lim, "Multiuser superposition transmission (MUST) for LTE-A systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.
[9] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.
[10] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 5, pp. 796–807, Jun. 1994.
[11] D. Tse and P. Viswanath, *Fundamentals Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[12] 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. 77th IEEE VTC-Spring*, Dresden, Germany, Jun. 2013, pp. 1–5.
[13] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, London, U.K., Sep. 2013, pp. 611–615.

[14] A. Benjebbovu, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, "System-level performance of downlink NOMA for future LTE enhancements," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2013, pp. 66–70.

[15] A. Benjebbour, K. Saito, A. Li, Y. Kishiyama, and T. Nakamura, "Non-orthogonal multiple access (NOMA): Concept, performance evaluation and experimental trials," in *Proc. Int. Conf. Wireless Netw. Mobile Commun. (WINCOM)*, Oct. 2015, pp. 1–6.

[16] 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.

[17] J. Choi, "Non-orthogonal multiple access in downlink coordinated two-point systems," *IEEE Commun. Lett.*, vol. 18, no. 2, pp. 313–316, Feb. 2014.

[18] 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.

[19] Q. Sun, S. Han, C. L. I, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 405–408, Aug. 2015.

[20] 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.

[21] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[22] Z. Yang, Z. Ding, P. Fan, and Z. Ma, "Outage performance for dynamic power allocation in hybrid non-orthogonal multiple access systems," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1695–1698, Aug. 2016.

[23] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7244–7257, Nov. 2016.

[24] 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.

[25] J. Cui, Z. Ding, and P. Fan, "A novel power allocation scheme under outage constraints in NOMA systems," *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1226–1230, Sep. 2016.

[26] S. N. Datta and S. Kalyanasundaram, "Optimal power allocation and user selection in non-orthogonal multiple access systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2016, pp. 1–6.

[27] L. Lei, D. Yuan, C. K. Ho, and S. Sun, "Joint optimization of power and channel allocation with non-orthogonal multiple access for 5G cellular systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[28] 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.

[29] B. Di, L. Song, and Y. Li, "Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7686–7698, Nov. 2016.

[30] S. Zhang, B. Di, L. Song, and Y. Li, "Radio resource allocation for non-orthogonal multiple access (NOMA) relay network using matching game," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[31] X. Li, C. Li, and Y. Jin, "Dynamic resource allocation for transmit power minimization in OFDM-based NOMA systems," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2558–2561, Dec. 2016.

[32] Z. Wei, D. W. K. Ng, and J. Yuan. (2016). "Power-efficient resource allocation for MC-NOMA with statistical channel state information." [Online]. Available: https://arxiv.org/abs/1607.01116

[33] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1077–1091, Mar. 2017.

[34] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Found. Trends Netw.*, vol. 1, no. 1, pp. 1–144, 2006.

[35] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. San Rafael, CA, USA: Morgan & Claypool Publishers, 2010.

[36] V. K. N. Lau and Y. Cui, "Delay-optimal power and subcarrier allocation for OFDMA systems via stochastic approximation," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 227–233, Jan. 2010.

[37] X. Wei and M. J. Neely, "Power-aware wireless file downloading: A Lyapunov indexing approach to a constrained restless bandit problem," *IEEE/ACM Trans. Network.*, vol. 24, no. 4, pp. 2264–2277, Aug. 2016.

[38] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. Cambridge, MA, USA: MIT Press, 2009.

Wei Bao (S'10–M'16) received the B.E. degree in communications engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2009, the M.A.Sc. degree in electrical and computer engineering from The University of British Columbia, Vancouver, Canada, in 2011, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, Canada, in 2016. He is currently a Lecturer with the School of Information Technologies, The University of Sydney, Sydney, Australia. His research covers the area of network science, with particular emphasis on 5G systems, Internet of Things, and mobile computing. He received the Best Paper Awards from the ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems in 2013 and the IEEE International Symposium on Network Computing and Applications in 2016.

He (Henry) Chen (S'10–M'16) received the Ph.D. degree in electrical engineering (telecommunications) from The University of Sydney, Sydney, Australia, in 2015. He is currently a Post-Doctoral Research Fellow with the School of Electrical and Information Engineering, The University of Sydney. His current research interests are in the field of wireless communications, with a particular focus on ultra-reliable low latency communications, non-orthogonal multiple access, millimeter wave wireless communications, and wireless energy harvesting and transfer.

Dr. Chen received the Outstanding Bachelor Thesis of Shandong University, the Outstanding Master Thesis of Shandong Province, and the Chinese Government Award for Outstanding Self-Financed Students Abroad. He served as a Guest Editor of the 2016 special issue of millimeter wave wireless communications and networks in MDPI Sensors, and an Organizer and a Co-Chair of GLOBECOM 2016 workshop on wireless energy harvesting communication networks. He has been serving as a regular reviewer for over 10 prestigious IEEE journals and the TPC members of several IEEE flagship conferences, such as GLOBECOM and ICC.

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

His current research interests are in the area of wireless communications, with a particular focus on MIMO, millimeter wave communications, machine to machine communications, coding techniques, and cooperative communications. He holds a number of patents granted and pending in these fields. He is currently an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and an Executive Editor of the *European Transactions on Telecommunications*. He received the Best Paper Award from the IEEE International Conference on Communications 2014 and the IEEE Wireless Days Conferences 2014.

Branka Vucetic (SM'00–F'03) is currently an ARC Laureate Fellow and a Professor of telecommunications, the Director of the Center of Excellence in Telecommunications, The University of Sydney. During her career she has held research and academic positions in Yugoslavia, Australia, U.K., and China. Her research interests include coding, communication theory, and signal processing and their applications in wireless networks and industrial Internet of Things.

She has co-authored four books and over four hundred papers in telecommunications journals and conference proceedings. She is a Fellow of the Australian Academy of Technological Sciences and Engineering.