Reconsidering Design of Multi-Antenna NOMA Systems With Limited Feedback

Zhiyao Tang, Liang Sun, *Member, IEEE*, Lu Cao, Shutong Qi, and Yong Feng, *Member, IEEE*

Abstract—We provide in this paper a comprehensive solution to the design, performance analysis, and optimization of a multi-antenna non-orthogonal multiple access (NOMA) system for multiuser downlink communications under a general limited channel state information (CSI) feedback framework for frequency division duplex mode. We design a general framework including user clustering, joint power and bits allocation, CSI quantization and feedback, signal superposition coding, transmit beamforming, and successive interference cancellation at receivers. Then, we conduct a mathematically strict performance analysis of the considered system, and obtain a closed-form lower bound on the ergodic rate of each user in terms of transmit power, CSI quantization accuracy and channel conditions. For exploiting the potentials of multiple-antenna techniques in NOMA systems, we jointly optimize two key parameters, i.e., transmit power and the number of feedback bits allocated to each user, and propose low-complexity closed-form solutions. Moreover, through asymptotic analysis, we reveal the interactions between the main system parameters and their impacts on the joint power and feedback bits allocation result, and hence show some guidelines on the system design. Finally, numerical results validate the correctness of our theoretical analysis and demonstrate the advantages of the proposed algorithms over the most related state of the art.

Index Terms—Non-orthogonal multiple access, multiple-antenna techniques, limited feedback, random vector quantization, performance analysis.

I. INTRODUCTION

NON-ORTHOGONAL multiple access (NOMA) has recently attracted considerable attentions as a promising technique for emerging services such as Internet of Things (IoT), the fifth generation (5G) mobile networks due to its potential in supporting massive connectivity and achieving high spectral efficiency [1], [2]. By combining superposition coding at the transmitters with successive interference cancellation (SIC) at the receivers, all the users in a NOMA system are able to decode their desired signal even though

they share the same frequency, time and space resources. The spectral efficiency improvement of NOMA technique over orthogonal multiple access (OMA) has been well investigated for single-antenna systems [3]–[8]. The spatial multiplexing property of multi-antenna systems can be naturally combined with power-domain NOMA to further enhance the spectral efficiency [9]–[21]. NOMA was also regarded as a spectrum sharing technique in [22].

It was noted in [9] that all existing works on multi-antenna NOMA systems can be classified into two categories: 1) there is a cluster of multiple users supported by each channel spatial dimension (or by a beamforming (BF) vector) in average that perform NOMA independently with the other clusters [9]–[15]; and 2) there is only one user supported by each channel spatial dimension in average[1] [16], [17]. For the first category, [2], [9], [11] proposed BF-based NOMA schemes, where joint design of user clustering, power allocation and BF were proposed to reduce the inter-cluster and intra-cluster interference. [12] purposed multi-antenna NOMA scheme with user grouping and clustering based on a long-term feedback along with user scheduling based on a short-term feedback. A closed-form hybrid-NOMA precoding algorithm was obtained in [14] based on the concept of quasi-degradation introduced in [13]. [15] designed linear precoders for multi-cell multi-antenna NOMA systems to maximize the total throughput. Obviously, the second category does not fully explore the spatial multiplexing capability of multi-antenna systems.

The quality of CSI available at transmitter (CSIT) and receiver of a multi-antenna system plays an important role and determines the performance of any multiple-antenna technique. In practice, the CSI acquisition methods can be mainly classified into the following two kinds. In time duplex division (TDD) systems, by employing the channel reciprocity of downlink and uplink channels, the transmitter can obtain the downlink CSI through channel estimation during the uplink training. While in frequency division duplex (FDD) systems, the downlink CSI is usually first estimated and quantized using a pre-determined codebook at each receiver, and then conveyed to transmitter via feedback channel. Specifically, the NOMA schemes in [9], [16], [17] were designed based on the kind of CSI acquisition in TDD mode, whilst the NOMA schemes in [18]–[20] were designed based on the kind of CSI acquisition in FDD mode. Particularly, the work in [21] had tried to employ a unified channel model to embrace both TDD

---

[1]The transmission of this category is the same as that in OMA systems with SIC receiver employed by some users.

and FDD modes. For both practical TDD and FDD modes, there will be residual inter-cluster and intra-cluster interference caused by the imperfect CSIT.

To the best of our knowledge, there have been very limited previous works on the transmit-receive design, performance analysis and optimization for the multi-antenna NOMA systems with quantized CSIT through limited feedback, with the very few exceptions in [18]–[21]. Specifically, employing ZFBF, [20] proposed a dynamic user scheduling and clustering strategy and optimized the net throughout for the delay-intolerant systems. Also based on ZFBF, [21] tried to propose a similar multi-antenna NOMA scheme to that in [20] with a unified imperfect CSIT model applicable to both FDD and TDD systems. The authors in [21] had also tried to jointly optimize the transmit power and feedback bits of each users. However, [21] did not use the well recognized practical CSI model for FDD systems as in [19], [20], [23], [24]. One can see the model used by [21] actually assumed some terms in the practical CSI model that should vary at the same speed of channel fading to be constant,[2] which is impractical for FDD systems. Moreover, we find it an essentially very hard problem to obtain the *exact* analytical performance of any multi-antenna NOMA scheme based on the linear ZFBF at the transmitter for general system settings (e.g. [19], [21]). The analytical performance of both schemes in [19], [21] were actually obtained by ignoring the statistical dependence between the useful signal term and the interference terms in each user's received signal. However, it is very difficult to analyze the impact of the aforementioned ignorance of the statistical dependence on the resulting analytical results.

In this paper, we provide a comprehensive study on the performance of multi-antenna NOMA and the joint power and feedback bits allocation optimization under the considered system. Our main contributions are summarized as follows:

- Motivated by the limitations with the previous works, we reconsider the design of the multi-antenna NOMA systems in FDD mode including user clustering, joint optimization of power and feedback bits allocation, limited CSI feedback, transmit BF and SIC at receivers. In this paper, we propose a low-complexity dynamic user clustering based on the large-scale fading of each user only, which can reduce the complexity of power allocation optimization and at the same time takes into consideration the user fairness.
- For a given power and feedback bits allocation, we provide a mathematically strict statistical analysis on the ergodic rate of each user taking into consideration the statistical dependence between the useful signal term and the interference terms in each user's received signal, which has never been done before. Our analysis based on the well recognized limited feedback framework results in a closed-form lower bound on the ergodic rate of each user without assuming any extreme for system parameters,

a result that, to the best of the authors knowledge, has not been previously presented in the literatures.
- For a given power allocation result to all users, we obtain a closed-form expression for the optimal feedback bits allocation (without integer constraint) to minimize an upper bound on the ESR loss due to quantized CSI feedback. Then, we devise a low-complexity dynamic programming algorithm to find the optimal practical bits allocation solution.
- For the transmit power allocation, we first follow the intuitive power allocation scheme in [21] to employ equal power allocation among the users within each cluster. Then, we observe that the ergodic rate of the users other than the nearest user (to the BS) in each cluster is relatively much smaller than that of the nearest user, and the ESR of these users within each cluster is also very limited compared with that of the nearest user. Thus, we seek for the solution by maximizing the ESR of the nearest users of all clusters, which is generally a non-convex optimization problem. Then, based on the closed-form solution of the feedback bits allocation obtained above, we propose a closed-form sub-optimal solution to the power allocation among the different clusters. Finally, a low-complexity scheme of joint optimization of the transmit power and feedback bits allocation is obtained.
- Through asymptotic analysis of the joint power and feedback bits allocation solution, we observe several key insights. Specially, for the scenario with high CSI quantization accuracy (or equivalently large total bandwidth of feedback channels B) and finite total transmit power P, our power allocation solution tends to be the well known water-filling type solution for the system with perfect CSIT. For the scenario in the high power region with finite B, our power allocation solution tends to be equal power allocation among all users. Moreover, in the high power region with large enough B, the optimal feedback bits allocated to the nearest user in each cluster is scaled approximately linearly increasing as $\log_2(P)$, whilst the corresponding one for the user other than the nearest user is scaled linearly approximately decreasing as $\log_2(P)$. For the scenario with finite B and large enough P, the number of feedback bits allocated to the user other than the nearest user in each cluster reduces to zero, whilst the one corresponding to the nearest user of each cluster approximately converges to a finite number that can be given in closed-form without integer constraint.

Notations: $\mathcal{C}$ and $\mathcal{N}$ denote the sets of complex numbers and natural numbers respectively. $\mathbb{E}_X\{\cdot\}$ represents expectation with respect to random variable X. $\|\cdot\|$ denotes the L_2-norm of a vector, $|\cdot|$ denotes the absolute value of a scalar.

II. SYSTEM AND SIGNAL MODELS

We consider downlink communications in a single-cell cellular network operating in FDD mode for delay-tolerant traffics, where the coding block of each user's information can be sufficiently long to cover multiple channel coherent time periods such that the ergodic rate can be used as the

---

[2]Specifically, the CSI model in [21] replaced the time-varying terms $\sin^2 \theta_{n,k}$ and $\cos^2 \theta_{n,k}$ of the channel model in [19], [20], [23], [24] (see (1) in this paper) by the constant terms $2^{-\frac{B_{n,k}}{M-1}}$ and $1 - 2^{-\frac{B_{n,k}}{M-1}}$ respectively.

performance metric. This scenario has been widely considered in the literature of NOMA [3], [7], [11], [21], [22]. There is a BS equipped with M antennas and a large number of users with single-antenna each. The BS simultaneously broadcasts information to multiple scheduled users by BF at each channel use. Power-domain NOMA technique is concurrently employed to further improve the spectral efficiency, where all NK scheduled users are grouped into N ($N \geq 1$) clusters with K ($K \geq 2$) users within each cluster being simultaneously supported by NOMA[3] [2]. Our user clustering method will be introduced in the following. We denote the M-dimensional channel vector from the BS to the k-th user in the n-th cluster (denoted as user (n, k) afterward) as $d_{n,k}^{-\frac{\alpha}{2}} \mathbf{h}_{n,k}$, where $\mathbf{h}_{n,k} \in \mathcal{C}^{1 \times M}$ and $d_{n,k}$ are respectively the fast-fading channel and the distance from the BS to user (n, k). All channels are assumed to undergo block flat-fading with a path-loss governed by the exponential coefficient α.

A. Channel and Feedback Models

Since we consider the system in FDD mode, to focus on the effect of limited feedback, we follow many previous works (e.g. [18]–[20]) to assume each user can perfectly estimate the downlink CSI, and then feeds back the quantized CSI to the BS through an error-free but limited-rate feedback channel. The total bandwidth of the feedback channels of all users is constrained to be B bits.

As [21], the CSI obtained by user (n, k) is quantized using a codebook $\mathbb{C}_{n,k}$, which consists of $2^{B_{n,k}}$ codewords. Here $B_{n,k}$ is the number of feedback bits allocated to user (n, k). Since the optimal CSI quantization strategy is unknown in general and is out of the scope of this work, we employ the same random quantization codebooks as those in [23], [24]. The rule for user (n, k) to quantize the channel direction information (CDI) (i.e., $\tilde{\mathbf{h}}_{n,k} = \frac{\mathbf{h}_{n,k}}{||\mathbf{h}_{n,k}||}$) is given by $\hat{\mathbf{h}}_{n,k} = \arg \max_{\mathbf{c} \in \mathcal{C}_{n,k}} \left| \tilde{\mathbf{h}}_{n,k} \mathbf{c}^H \right|$. Then, $\tilde{\mathbf{h}}_{n,k}$ can be decomposed as [23]

$$\tilde{\mathbf{h}}_{n,k} = \cos \theta_{n,k} \hat{\mathbf{h}}_{n,k} + \sin \theta_{n,k} \tilde{\mathbf{e}}_{n,k}, \qquad (1)$$

where $\theta_{n,k} = \angle(\tilde{\mathbf{h}}_{n,k}, \hat{\mathbf{h}}_{n,k})$ and $\tilde{\mathbf{e}}_{n,k}$ is the normalized quantization error vector that is isotropically distributed in the nullspace of $\hat{\mathbf{h}}_{n,k}$. Note that, due to very limited bandwidth of feedback channels, only the index of quantized CDI in the codebook of each user is fed back. For the transmitter to obtain the large-scale fading (or $d_{n,k}^{-\alpha}$), each user measures "reference signal receiving power" through downlink cell-specific reference signals (for long term evolution (LTE) systems) or CSI-reference signals (for LTE-advanced systems) [25] and then feeds back mean received signal power to the transmitter. Since the required frequency to feed back this information is much lower than that of fast fading (i.e., quantized CDI), we ignore its cost and follow many previous related papers to assume path loss of each user is known by the transmitter.

B. User Clustering

Most of the existing user clustering methods for multi-antenna NOMA systems are according to fast-fading channels. However, this kind of methods carries out exhaustive search resulting in high implementation complexity, and additionally requires feedback of accurate enough channel quality information (e.g., channel gain or SINR) for user clustering and determining the order of SIC within each cluster, which costs much more bandwidth of feedback channels for them to work correctly. In addition, user clustering also depends on the power allocation to each user in general, which makes the kind of the methods based on instantaneous CSI even more complicated. In spite of all existing methods, considering delay-tolerant traffics and only very limited CSIT of each user,[4] we propose a low-complexity clustering strategy based on the large-scale fading of each user only. Considering the user fairness, our scheme first randomly selects NK users each time, and then forms N user clusters with K users within each following the criteria that

$$\frac{d_{n,i}^{-\alpha}}{\sigma_{n,i}^2} \leq \frac{d_{m,i}^{-\alpha}}{\sigma_{m,i}^2} \quad \text{for } 1 \leq m < n \leq N, \ \forall 1 \leq i \leq K; \quad (2)$$

$$\frac{d_{n,i}^{-\alpha}}{\sigma_{n,i}^2} \leq \frac{d_{n,j}^{-\alpha}}{\sigma_{n,j}^2} \quad \text{for } 1 \leq j \leq i \leq K, \ \forall 1 \leq n \leq N, \quad (3)$$

where $\sigma_{n,k}^2$ is the variance of the additive white Gaussian noise (AWGN) at user (n, k). It will be shown in Section IV-B after Remark 7 that, the criteria working with our proposed joint power and feedback bits allocations maximize the system ESR.

C. Zero-Forcing Beamforming and Successive Interference Cancellation

To mitigate MU interference, we employ the widely used simple linear ZFBF at the BS [19]–[21], [23], [24], where the BF vector $\mathbf{w}_n$ of cluster n satisfies $\hat{\mathbf{h}}_{m,k} \mathbf{w}_n = 0$ $\forall m \neq n$ and $\forall k$. Following the method in [21], for cluster n we construct a complementary matrix $\bar{\mathbf{H}}_n = \left[\hat{\mathbf{h}}_{1,1}^H, \ldots, \hat{\mathbf{h}}_{1,K}^H, \ldots, \hat{\mathbf{h}}_{n-1,K}^H, \hat{\mathbf{h}}_{n+1,1}^H, \ldots, \hat{\mathbf{h}}_{N-1,K}^H, \hat{\mathbf{h}}_{N,1}^H, \ldots, \hat{\mathbf{h}}_{N,K}^H \right]^H \in \mathcal{C}^{(N-1)K \times M}$. Let the singular-value decomposition (SVD) of $\bar{\mathbf{H}}_n$ be $\bar{\mathbf{H}}_n = \bar{\mathbf{U}}_n \bar{\mathbf{\Sigma}}_n [\bar{\mathbf{V}}_n^{(1)}, \bar{\mathbf{V}}_n^{(0)}]$, where the columns of matrix $\bar{\mathbf{V}}_n^{(0)} \in \mathcal{C}^{M \times (M - \bar{r}_n)}$ are the right-singular vectors corresponding to the zero singular-values. Here $\bar{r}_n \triangleq \text{rank}(\bar{\mathbf{H}}_n) = (N-1)K$ with probability 1 when the channel vector of each user is continuously distributed. Then, $\mathbf{w}_n$ can be obtained as $\mathbf{w}_n = \bar{\mathbf{V}}_n^{(0)} \mathbf{p}_n$ [21], where $\mathbf{p}_n$ is uniformly distributed on the surface of the $(M - \bar{r}_n)$-dimensional unit sphere. It is easy to see the constraint $M > (N - 1)K$ needs to be satisfied for ZFBF to work properly. Then, with the decomposition of channel in (1), the received signal at user (n, k) can be

---

[3]For the convenience of presentation, we have assumed the number of users of each cluster to be equal to K. It is easy to see, with very simple generalization, our NOMA framework can simultaneously support arbitrary number of users (with some constraint by M), and can be applicable to the scenarios with different number of users in each cluster.

[4]Recall that only quantized CDI and the large-scale fading (path-loss) are available at the transmitter. Thus, user clustering method based on full CSIT of fast-fading channel at the transmitter is impossible.

written as

$$
\begin{aligned}
y_{n,k} &= d_{n,k}^{-\frac{\alpha}{2}} \mathbf{h}_{n,k} \sum_{i=1}^{N} \mathbf{w}_i s_i + n_{n,k} \\
&= d_{n,k}^{-\frac{\alpha}{2}} ||\mathbf{h}_{n,k}|| \sin\theta_{n,k} \sum_{i=1,i\neq n}^{N} \tilde{\mathbf{e}}_{n,k} \mathbf{w}_i s_i \\
&\quad + d_{n,k}^{-\frac{\alpha}{2}} \mathbf{h}_{n,k} \mathbf{w}_n s_n + n_{n,k},
\end{aligned}
$$

where $s_i = \sum_{j=1}^{K} \sqrt{P_{i,j}} s_{i,j}$ is the superposition coded signals for all users in cluster i with $s_{i,j}$ and $P_{i,j}$ being the signal of user (i,j) and the corresponding power, and $n_{n,k}$ being the AWGN. In general, $P_{i,j}$ should be properly allocated to enable successful SIC and ensure certain level of ESR performance of each user and also the fairness among the users within a cluster.

Recall that it is assumed the coding block of the considered systems can be sufficiently long. Thus, according to information theory, if the transmission rate of user (n,k) does not exceed the ergodic rate of user (n,k) supported by the multiuser fading channels, the decoding error with SIC at user (n,k) can be arbitrarily small as the coding block length of each user goes large. The above context was also assumed in [21], although it was not explicitly explained therein. Moreover, for the same reason as noted in [20], [21], when the power is properly allocated among the signals of the users within each cluster, user (n,k) can always successfully decode the user (n,j)'s signal for $\forall j > k$ with the user ordering given by (3), if user (n,k) can decode its own signal.[5] As a result, before decoding its own signal, user (n,k) can cancel the interference from user (n,j) for $j > k$ in the received signal. Thus, we can assume perfect SIC can be performed at each user as [20], [21]. Then, after SIC, the SINR at user (n,k) for $k=1$ and $k>1$ are given respectively by (4) and (5) at the bottom of this page. The SIC may not be perfect in practice due to limited capability of users. However, the study of the impact of imperfect SIC is out of the scope of this paper and will be considered in the future work.

The power allocation and feedback bits allocation have significant impact on the performance of the NOMA systems. Before optimizing the power and feedback bits allocation, we first analyze the ergodic rate of the considered system.

[5]It is easy to see that our proposed method in Subsection IV-B that allocates equal power to the users within each cluster satisfies this requirement.

III. PERFORMANCE ANALYSIS

In this section, we focus on obtaining the analytical ergodic rate of each user given power and bits allocation and without assuming any extreme for system parameters. All fast-fading channels are assumed to follow circularly symmetric complex Gaussian distribution with zero mean and unit variance, i.e., Rayleigh fading. We note that, since the user ordering given by (2) and (3) does not depend on the fast fading, the distribution of each fast-fading term of different scheduled users is the same. The ergodic rate of user (n,k) is given by $R_{n,k} = \mathbb{E}[\log_2(1+\gamma_{n,k})]$ with $\gamma_{n,k}$ given by (4) or (5) shown at the bottom of this page. First, we note that, due to the statistical dependence between useful signal term and interference terms at each user, it is very difficult if not impossible to obtain the exact analytical result of $R_{n,k}$. Thus, we will turn to obtaining an analytical result as accurate as possible. Specifically, we can obtain a lower bound of $R_{n,k}$ in the following theorem.

Theorem 1: $R_{n,k}$ $\forall n, k$ can be lower-bounded as $R_{n,k} \geq R_{n,k}^{LB1}$ with $R_{n,k}^{LB1}$ given by

$$
R_{n,1}^{LB1} = \Theta_{n,k}(\alpha_{n,1} P_{n,1}, S_{n,1}^{(3)}(M-1)^{-1}\delta_{n,1}) - \log_2(e)
$$
$$
\times \exp\left(\frac{M-1}{S_{n,1}^{(3)}\delta_{n,1}}\right) \sum_{q=1}^{M-1} E_q\left(\frac{M-1}{S_{n,1}^{(3)}\delta_{n,1}}\right), \quad (6)
$$
$$
R_{n,k}^{LB1} = \Theta_{n,k}\left(S_{n,k}^{(1)}, S_{n,k}^{(3)}(M-1)^{-1}\delta_{n,k}\right)
$$
$$
\quad - \Theta_{n,k}\left(S_{n,k}^{(2)}, S_{n,k}^{(3)}(M-1)^{-1}\delta_{n,k}\right), \quad \text{for } k > 1, \quad (7)
$$

where $S_{n,k}^{(1)} = \frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2} \sum_{j=1}^{k} P_{n,j}$, $S_{n,k}^{(2)} = \frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2} \sum_{j=1}^{k-1} P_{n,j}$, $S_{n,k}^{(3)} = \frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2} \sum_{i=1,i\neq n}^{N} \sum_{j=1}^{K} P_{i,j}$, $\delta_{n,k} = 2^{-\frac{B_{n,k}}{M-1}}$, and $E_q(x) = \int_1^{\infty} \frac{e^{-xt}}{t^q} dt$ is generalized exponential integral. $\Theta_{n,k}(a,b)$ is given by

$$
\begin{aligned}
\Theta_{n,k}(a,b) = \log_2(e) \Bigg[& \frac{(-1)^M(M-1)}{b^{M-1}} I_1(a) \\
& + \sum_{p=1}^{M-1} \sum_{q=1}^{M-p} \frac{(-1)^{p+q+1}(M-1)}{(q-1)!b^p} I_2(a,b,p,q) + \\
& \sum_{p=1}^{M-1} \sum_{q=2}^{M-p} \sum_{s=0}^{q-2} \frac{(-1)^{p+s-1}(q-s-2)!(M-1)}{(q-1)!b^p} I_3(a,b,p,s) \Bigg]
\end{aligned} \quad (8)
$$

$$
\gamma_{n,1} = \frac{d_{n,1}^{-\alpha}|\mathbf{h}_{n,1}\mathbf{w}_n|^2 P_{n,1}}{\underbrace{d_{n,1}^{-\alpha}\sin^2\theta_{n,1}||\mathbf{h}_{n,1}||^2 \sum_{i=1,i\neq n}^{N} |\tilde{\mathbf{e}}_{n,1}\mathbf{w}_i|^2 \sum_{l=1}^{K} P_{i,l}}_{\text{Inter-cluster interference}} + \sigma_{n,1}^2} \quad (4)
$$

$$
\gamma_{n,k} = \frac{d_{n,k}^{-\alpha}|\mathbf{h}_{n,k}\mathbf{w}_n|^2 P_{n,k}}{\underbrace{d_{n,k}^{-\alpha}|\mathbf{h}_{n,k}\mathbf{w}_n|^2 \sum_{j=1}^{k-1} P_{n,j}}_{\text{Intra-cluster interference}} + \underbrace{d_{n,k}^{-\alpha}\sin^2\theta_{n,k}||\mathbf{h}_{n,k}||^2 \sum_{i=1,i\neq n}^{N} |\tilde{\mathbf{e}}_{n,k}\mathbf{w}_i|^2 \sum_{l=1}^{K} P_{i,l}}_{\text{Inter-cluster interference}} + \sigma_{n,k}^2} \quad (5)
$$

with $I_1(a)$, $I_2(a,b,p,q)$ and $I_3(a,b,p,s)$ given respectively as

$$I_1(a) = \sum_{t=0}^{M-2} \binom{M-2}{t} \frac{(-1)^t}{a^{t+1}} \tilde{\Psi}\left(-M-1-t, \frac{1}{a}\right),$$

$$I_2(a,b,p,q) = \sum_{r=0}^{M-2} \sum_{t=0}^{p-1+r} \binom{M-2}{r} \binom{p-1+r}{t}$$
$$\times \frac{(-1)^{p-1-t} b^{p-1+r-t}}{a^{r+1}} \Psi\left(q-4-t, \frac{1}{a+b}, \frac{1}{b}\right),$$

$$I_3(a,b,p,s) = \sum_{t=0}^{M-2} \binom{M-2}{t} \frac{(-1)^t a^{p-1}}{b^{s-1}(p+t)}$$
$$\times {}_2F_1\left(s-1, p+t; p+t+1; -\frac{a}{b}\right),$$

where ${}_2F_1(a,b;c;z)$ is a Gauss-hypergeometric function [26, 9.14.2] and $\Psi(n,u,v)$ is given by

$$\Psi(n,u,v) \triangleq$$
$$\begin{cases}
\sum_{k=0}^{n} \frac{(-1)^k n! \left[v^{n-k} e^v E_i(-v) - u^{n-k} e^u E_i(-u)\right]}{(n-k)!} \\
\quad - \sum_{k=0}^{n-1} \frac{(-1)^k n! (v^{n-k} - u^{n-k})}{(n-k)(n-k)!} - (-1)^n n! \ln\left(\frac{v}{u}\right), & n > 0 \\
e^v E_i(-v) - e^u E_i(-u) - \ln\left(\frac{v}{u}\right), & n = 0 \\
E_i(-v) E_i(v) - E_i(-u) E_i(u) + \frac{E_i^2(-u) - E_i^2(-v)}{2} \\
\quad + \frac{1}{2} \sum_{m=0}^{\infty} \sum_{l=0}^{2m} \frac{2(v^l e^{-v} - u^l e^{-u})}{(2m+1)^2 l!}, & n = -1 \\
\sum_{k=1}^{-n-1} \frac{(-n-k-1)!}{(-n-1)!} \left[u^{n+k} e^u E_i(-u) - v^{n+k} e^v E_i(-v)\right] \\
\quad + \frac{E_i(-v) E_i(v) - E_i(-u) E_i(u)}{(-n-1)!} + \\
\sum_{k=1}^{-n-1} \frac{(-n-k-1)!(v^{n+k} - u^{n+k})}{(-n-1)!(n+k)} + \frac{E_i^2(-u) - E_i^2(-v)}{2(-n-1)!} \\
\quad + \frac{1}{(-n-1)!} + \sum_{m=0}^{\infty} \sum_{l=0}^{2m} \frac{2(v^l e^{-v} - u^l e^{-u})}{(2m+1)^2 l!}, & n \le -2.
\end{cases}$$
(9)

$E_i(x) = -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt$ is the exponential integral. For $n < 0$ with $v \to +\infty$, $\Psi(n,u,v)$ becomes

$$\tilde{\Psi}(n,u) \triangleq \Psi(n,u,v)\big|_{v \to +\infty}$$
$$= \frac{e^u E_i(-u)}{(-n-1)!} \sum_{k=1}^{-n-1} (-n-k-1)! u^{n+k} - \frac{E_i(-u) E_i(u)}{(-n-1)!}$$
$$- \sum_{k=1}^{-n-1} \frac{(-n-k-1)! \, u^{n+k}}{(-n-1)!(n+k)} + \frac{E_i^2(-u)}{2(-n-1)!}$$
$$- \frac{1}{(-n-1)!} \sum_{m=0}^{\infty} \sum_{l=0}^{2m} \frac{2u^l e^{-u}}{(2m+1)^2 l!}.$$
(10)

Proof: See Appendix A. □

Remark 1: As it was also observed by the authors in [20], the exact analytical performance of the multi-antenna NOMA systems based on linear ZFBF with limited CSI feeback model here is very difficult to obtain. We believe the actual analytical result (if it could be obtained) is *at least as the same complicated as* the bound obtained in *Theorem* 1, if not more complicated. Thus, it is natural that any accurate analytical lower or upper bound should be very complicated. In spite of this, the theoretical result obtained in Theorem 1 is able to avoid the time-consuming computer simulations in evaluating the system performance.

Remark 2: It can be observed from (27) in *Appendix* A, the channel model with limited feedback used in [21] in fact equivalently further employs Jensen's inequality on the random term $\sin^2 \theta_{n,k}$ in (27), thus results in lower ergodic rate performance than that of any practical system. This point will also be verified by the numerical result of Fig. 2 in Section V.

IV. JOINT OPTIMIZATION OF FEEDBACK BITS AND POWER ALLOCATION

The signal strength can be enhanced and the MU interference can be reduced by increasing the accuracy of CSI quantization. The transmit power allocation to each user also has a great impact on the ergodic rate performance. Moreover, it is obvious that the feedback bits allocation and the power allocation interact with each other. Therefore, it is essential to consider the joint optimization of the feedback bits and power allocation among the MUs for performance enhancement with the constraints on the total transmit power and the sum bandwidth of feedback channels of all users.

A. Feedback Design Given Power Allocation Among All Users

In this subsection, we first concentrate on optimizing the ESR by feedback bits allocation among all users given the power allocation among all users. The power allocation optimization will be considered in the next subsection. [21] proposed a indirect method that minimized the average sum power of inter-cluster interference of all users, which can not guarantee maximizing the ESR. In contrast to [21], we consider directly maximizing the ESR of all users or equivalently minimizing ESR loss caused by CSI quantization. The ergodic rate loss of user (n,k) due to CSI quantization is defined as $\Delta R_{n,k} \triangleq R_{n,k}^{ideal} - R_{n,k}$, where $R_{n,k}^{ideal} = \mathbb{E}\left[\log_2\left(\frac{1 + S_{n,k}^{(1)} |\mathbf{h}_{n,k} \mathbf{w}_n|^2}{1 + S_{n,k}^{(2)} |\mathbf{h}_{n,k} \mathbf{w}_n|^2}\right)\right]$ is the ergodic rate with perfect CSI. Since the actual analytical result of $R_{n,k}$ is unknown and the obtained analytical lower bound in Section III is too sophisticated for further processing, in the following theorem we first develop an upper bound on the ergodic rate loss with a relatively simpler form.

Theorem 2: $\Delta R_{n,k}$ $\forall n, k$ can be upper-bounded as $\Delta R_{n,k} \le \Delta R_{n,k}^{UB1}$, where

$$\Delta R_{n,k}^{UB1} = \log_2\left(1 + S_{n,k}^{(2)} + \Gamma\left(\frac{2M-1}{M-1}\right) 2^{-\frac{B_{n,k}}{M-1}} S_{n,k}^{(3)}\right)$$
$$- \log_2(e) e^{\frac{1}{S_{n,k}^{(2)}}} E_1\left(\frac{1}{S_{n,k}^{(2)}}\right).$$
(11)

Here $S_{n,k}^{(2)}$ and $S_{n,k}^{(3)}$ are the same as defined in *Theorem* 1.

Proof: See Appendix D. □

With *Theorem* 2, an upper bound on the ESR loss ΔR_{sum} can be obtained as

$$\Delta R_{sum,UB} = \sum_{n=1}^{N} \sum_{k=1}^{K} \Delta R_{n,k}^{UB1}.$$
(12)

Then, the feedback bits allocation optimization problem turns into the problem of minimizing $\Delta R_{sum,UB}$. Moreover, since the second term of (11) is irrelevant to $B_{n,k}$, the problem can be equivalently formulated as

$$\min_{\{B_{n,k}\}} \prod_{n=1}^{N} \prod_{k=1}^{K} \left(1 + S_{n,k}^{(2)} + \Gamma \left(\frac{2M-1}{M-1} \right) 2^{-\frac{B_{n,k}}{M-1}} S_{n,k}^{(3)} \right)$$

$$\text{s.t.} \sum_{n=1}^{N} \sum_{k=1}^{K} B_{n,k} \leq B, \tag{13}$$

$$B_{n,k} \in \mathcal{N}, \quad \forall n, k.$$

Without non-negative integer constraint, the solution to (13) is obtained in the following theorem.

Theorem 3: The solution to the problem (13) without non-negative integer constraint is given by

$$B_{n,k}^{\star} = \frac{B}{NK} + (M-1) \log_2 \left(\frac{S_{n,k}^{(3)}}{1 + S_{n,k}^{(2)}} \right)$$

$$+ \frac{(M-1)}{NK} \sum_{p=1}^{N} \sum_{q=1}^{K} \log_2 \left(\frac{1 + S_{p,q}^{(2)}}{S_{p,q}^{(3)}} \right). \tag{14}$$

Proof: See Appendix E. □

Remark 3: By employing the arithmetic-geometric mean inequality, we have

$$\prod_{n=1}^{N} \prod_{k=1}^{K} \left(1 + S_{n,k}^{(2)} + \Gamma \left(\frac{2M-1}{M-1} \right) 2^{-\frac{B_{n,k}}{M-1}} S_{n,k}^{(3)} \right)$$

$$\leq \left[\frac{1}{NK} \sum_{n=1}^{N} \sum_{k=1}^{K} \left(1 + S_{n,k}^{(2)} + \Gamma \left(\frac{2M-1}{M-1} \right) \right.\right.$$

$$\left.\left. \times 2^{-\frac{B_{n,k}}{M-1}} S_{n,k}^{(3)} \right) \right]^{NK}$$

$$= \left[1 + \frac{1}{NK} \left(\sum_{n=1}^{N} \sum_{k=1}^{K} S_{n,k}^{(2)} + \Gamma \left(\frac{2M-1}{M-1} \right) \right.\right.$$

$$\left.\left. \times \sum_{n=1}^{N} \sum_{k=1}^{K} 2^{-\frac{B_{n,k}}{M-1}} S_{n,k}^{(3)} \right) \right]^{NK}. \tag{15}$$

We notice the term $\sum_{n=1}^{N} \sum_{k=1}^{K} 2^{-\frac{B_{n,k}}{M-1}} S_{n,k}^{(3)}$ in the right hand side (RHS) of (15) is exactly the objective function of the optimal bits allocation problem formulated in [21, (24)]. Since the other terms in the RHS of (15) is irrelevant to $\{B_{n,k}\}$, this illustrates minimizing the objective function of [21] is equivalent to minimizing an upper bound of the objective function of our problem. Thus, our algorithm is superior to that in [21] in general. We will also show the superiority of our algorithm by the numerical results in Section V.

For some system parameter settings, (14) may give negative values (particularly for user (n,k) with $k > 1$ in each cluster). In this case, our algorithm temporally allocates 0 bit to these users. Then, the bits allocation algorithm given by *Theorem* 3 is performed again for the remaining users. This algorithm operates in the recurrent manner until the bits allocation for all remaining users satisfies $B_{n,k} \geq 0$. Without causing confusion, we still denote the resulting bits allocation of user

(n,k) after the above operation as $B_{n,k}^{\star}$. Moreover, there is integer constraint on each $B_{n,k}$. Thus, based on the coarse bits allocation results obtained above, we seek for the practical solution by *Algorithm* 1.

Algorithm 1 The Dynamic Programming Algorithm to Find Practical Feedback Bits Allocation

---

Input: $\mathbf{b}_{rd}$, $\mathbf{r}_v(i)$, B_{re}
Output: $\mathbf{b}_{int}$
1: $\mathbf{D}_p = \text{zeros}(NK + 1, B_{re} + 1)$;
2: $\mathbf{B}_{rc} = \text{zeros}(NK, NK + 1, B_{re} + 1)$;
3: **for** $i = 2$ to $NK + 1$ **do**
4:     **for** $j = 2$ to $B_{re} + 1$ **do**
5:         **if** $\mathbf{D}_p(i-1, j) > \mathbf{D}_p(i-1, j-1) + \mathbf{r}_v(i-1)$ **then**
6:             $\mathbf{B}_{rc}(:, i, j) = \mathbf{B}_{rc}(:, i-1, j)$;
7:             $\mathbf{D}_p(i, j) = \mathbf{D}_p(i-1, j)$;
8:         **else**
9:             $\mathbf{B}_{rc}(:, i, j) = \mathbf{B}_{rc}(:, i-1, j-1)$;
10:          $\mathbf{B}_{rc}(i, i, j) = 1$;
11:          $\mathbf{D}_p(i, j) = \mathbf{D}_p(i-1, j-1) + \mathbf{r}_v(i-1)$;
12: **return** $\mathbf{b}_{int} = \mathbf{b}_{rd} + \mathbf{B}_{rc}(:, NK + 1, B_{re} + 1)$;

First, let $\mathbf{b}_{int} \in \mathcal{N}^{NK \times 1}$ consist of the integer bits allocation to all users, where the i-th element $\mathbf{b}_{int}(i)$ is the number of bits allocated to user (n', k') with the relationship between i and (n', k') given by $n' = \text{MOD}(i - 1, N) + 1$ and $k' = \lceil i/N \rceil$. Here MOD is the modular operation. Similarly, let $\mathbf{b}_{rd} \in \mathcal{N}^{NK \times 1}$ denote the bits allocation obtained as $\mathbf{b}_{rd}(i) = \lfloor \{B_{n', k'}^{\star}\} \rfloor$. In addition, let $\mathbf{r}_v$ be $NK \times 1$ vector with the i-th element $\mathbf{r}_v(i) = \Delta R_{n', k'}^{UB1}(\mathbf{b}_{rd}(i)) - \Delta R_{n', k'}^{UB1}(\mathbf{b}_{rd}(i) + 1)$ denoting the gain of the ergodic rate from allocating one extra bit to user (n', k'), where $\Delta R_{n', k'}^{UB1}$ is given by (11). Let $B_{re} = B - \sum_{i=1}^{NK} \mathbf{b}_{rd}(i)$ denote the number of the remaining bits that need to be re-allocated. Then, the problem to find the practical optimal solution can be formulated as

$$\max_{\mathbf{b}_{int}} [\mathbf{b}_{int} - \mathbf{b}_{rd}]^T \mathbf{r}_v \tag{16}$$

$$\text{s.t.} \sum_{i=1}^{NK} [\mathbf{b}_{int}(i) - \mathbf{b}_{rd}(i)] \leq B_{re},$$

$$[\mathbf{b}_{int}(i) - \mathbf{b}_{rd}(i)] \in \{0, 1\}.$$

The problem in (16) is a classical 0/1 Knapsack Problem [27] and can be solved by the dynamic programming algorithm. The detailed steps of the algorithm is provided in *Algorithm* 1. Finally, we obtain the bits allocation result with integer constraint given by $B_{n,k}^{int\star} = \mathbf{b}_{int}(i)$ with $i = N(k-1) + n$.

Remark 4: The overall complexity of the dynamic programming algorithm in *Algorithm* 1 is mainly determined by the complexity of at each iteration and the number of iterations required. The computations within each iteration are just a few scalar additions. Moreover, the time complexity of the dynamic programming algorithm shown in *Algorithm* 1 is $O(NKB_{re})$, and it is easy to see $B_{re} \leq \lceil NK/2 \rceil$. Therefore, the time complexity of the algorithm is a polynomial of the scale of the system, i.e., the parameters N, K. Thus, this is a very low-complexity algorithm.

B. Power Allocation

The power allocation in NOMA systems is more complex than that in OMA systems. For the power allocation within one user cluster, according to the principle of NOMA, on the one hand the power allocation usually follows the criterion that $P_{n,1} \leq \cdots \leq P_{n,K}$ in order to facilitate SIC and guarantee certain level of quality of service (QoS) requirements of the users[6] with poor channel conditions [11]. On the other hand, in order to maximize the ESR, the power allocation should follow the criterion that $P_{n,1} \geq \cdots \geq P_{n,K}$. Therefore, [21] proposed equal power allocation within a cluster to satisfy the above two criterions simultaneously and to reduce the complexity of the algorithm. We will follow the same scheme, i.e., $P_{n,k} = P_n/K$, where P_n is the total power allocated to the users in the n-th cluster. For the power allocation among the clusters, recall that power allocation and bits allocation actually *interact with each other*. Thus, in contrast to the intuitive scheme in [21] which was based on a given feedback bits allocation (e.g., equal feedback bit allocation among all users in the numerical result of [21]), we propose to *jointly* optimize these two parameters by maximizing the system ESR under the premise of $P_{n,k} = P_n/K \ \forall n, k$.

For the tractability, we develop in Appendix F another analytical lower bound of $R_{n,k} \ \forall n, k$ as

$$R_{n,k} \geq R_{n,k}^{LB2} \triangleq \Delta R_{n,k}^{LB2} + \tilde{R}_{n,k}^{LB2} \tag{17}$$

with $\Delta R_{n,k}^{LB2} = \mathbb{E}\left[\log_2\left(\frac{1+\bar{S}_{n,k}^{(1)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2}{1+\bar{S}_{n,k}^{(1)}M^{-1}||\mathbf{h}_{n,k}||^2}\right)\right]$, which is a negative number independent with bits allocation and no smaller than the constant $-\log_2(e)\mathbf{C} \approx -0.8327$ for any system parameters.[7] where $\mathbf{C}$ is the Eulers constant [26, 8.367.2]. $\Delta R_{n,k}^{LB2}$ changes very slowly with P_n, thus can be identified P_n-independent for power allocation. In addition,

$$\tilde{R}_{n,k}^{LB2}$$
$$= \log_2\left(1+\frac{\frac{d_{n,k}^{-\alpha}P_n}{\sigma_{n,k}^2 K}}{\frac{M}{M-1}+\bar{S}_{n,k}^{(2)}+\Gamma\left(\frac{2M-1}{M-1}\right)2^{-\frac{B_{n,k}}{M-1}}\bar{S}_{n,k}^{(3)}}\right). \tag{18}$$

Here, $\bar{S}_{n,k}^{(1)} = \frac{d_{n,k}^{-\alpha}kP_n}{\sigma_{n,k}^2 K}$, $\bar{S}_{n,k}^{(2)} = \frac{d_{n,k}^{-\alpha}(k-1)P_n}{\sigma_{n,k}^2 K}$ and $\bar{S}_{n,k}^{(3)} = \frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2}\sum_{i=1,i\neq n}^N P_i = \frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2}(P-P_n)$ are obtained by substituting $P_{n,k} = P_n/K$ into $S_{n,k}^{(1)}$, $S_{n,k}^{(2)}$ and $S_{n,k}^{(3)}$ respectively. Specifically, $\bar{S}_{n,k}^{(2)} = 0$ for $k = 1$.

We find it still very complicated to take $\sum_{n=1}^N \sum_{k=1}^K \tilde{R}_{n,k}^{LB2}$ as the objective function to optimize power allocation. Thus, we will further simplify the problem. We can show, with *perfect* CSIT and $P_{n,k} = P_n/K$, $R_{n,k}$ for $k > 1$ can be

upper bound as

$$R_{n,k} = \log_2\left(1+\frac{\frac{d_{n,k}^{-\alpha}P_n}{\sigma_{n,k}^2 K}|\mathbf{h}_{n,k}\mathbf{w}_n|^2}{1+\frac{(k-1)d_{n,k}^{-\alpha}P_n}{\sigma_{n,k}^2 K}|\mathbf{h}_{n,k}\mathbf{w}_n|^2}\right)$$
$$< \log_2\left(1+\frac{\frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2 K}|\mathbf{h}_{n,k}\mathbf{w}_n|^2}{\frac{(k-1)d_{n,k}^{-\alpha}}{\sigma_{n,k}^2 K}|\mathbf{h}_{n,k}\mathbf{w}_n|^2}\right) = \log_2\left(1+\frac{1}{k-1}\right),$$

which, compared to $R_{n,1}$, is a very limited amount independent with power allocation and decreasing with the user index. Moreover, the ESR of these users in each cluster is bounded as $\sum_{k=2}^K R_{n,k} < \log_2(K)$. The corresponding rates with limited CSI feedback must be even smaller. Thus, to further consider the ergodic rate of all users within each cluster in optimization can achieve very little benefit of sum rate, but it take much more computation complexity to search for a local-optimal solution due to the non-convex property of the sum ergodic rate. In addition, the equal power allocation within each cluster already guarantees certain transmission rate for these users. Therefore, in the following we take the sum of the terms $\tilde{R}_{n,k}^{LB2}$ in (18) for $k = 1$ as the objective function, i.e., $\tilde{R}_{sum}^{LB2}(1) \triangleq \sum_{n=1}^N \tilde{R}_{n,1}^{LB2}$.

Let $P_n = \phi_n P$, where the coefficients ϕ_n ($n = 1, 2, \cdots, N$) satisfy $\sum_{i=1}^N \phi_n = 1$. Substituting the result of the bits allocation without integer constraint given by (14) into the expression of $\tilde{R}_{n,k}^{LB2}$, after some manipulations we have

$$\tilde{R}_{n,k}^{LB2}(\{\phi_n\}) = \log_2\left(1+\frac{d_{n,k}^{-\alpha}\phi_n P}{K\sigma_{n,k}^2 D_{n,k}}\right), \tag{19}$$

where $D_{n,k} = \frac{M}{(M-1)} + \frac{d_{n,k}^{-\alpha}(k-1)\phi_n P}{\sigma_{n,k}^2 K} + \Gamma\left(\frac{2M-1}{M-1}\right) \times 2^{-\frac{B}{NK(M-1)}}\left(\prod_{p=1}^N \prod_{q=1}^K \left(\frac{d_{p,q}^{-\alpha}}{\sigma_{p,q}^2}\right)^{\frac{1}{NK}}\right)\left(1+\frac{d_{n,k}^{-\alpha}(k-1)\phi_n P}{\sigma_{n,k}^2 K}\right)$
$\left(\prod_{p=1}^N (1-\phi_p)^{\frac{1}{N}}\right)\left(\prod_{p=1}^N \prod_{q=1}^K \left(\frac{1}{P}+\frac{d_{p,q}^{-\alpha}(q-1)\phi_p}{\sigma_{p,q}^2 K}\right)^{-\frac{1}{NK}}\right)$.

Then, our power allocation problem becomes

$$\min_{\{\phi_n\}} \ -\tilde{R}_{sum}^{LB2}(1) = -\sum_{n=1}^N \tilde{R}_{n,1}^{LB2}(\{\phi_n\}) \tag{20}$$
$$\text{s.t. } \sum_{n=1}^N \phi_n \leq 1, \ \phi_n \geq 0.$$

It is not difficult to check that the problem in (20) is still a non-convex optimization problem. Thus, we will seek for a sub-optimal (local-optimal) solution, which is given by the following theorem.

Theorem 4: Given N and when transmit power P is large enough to support the N clusters of NOMA users, a sub-optimal solution to the problem (20) is given by

$$\phi_n^\star = \frac{1}{N} - \frac{K}{N}\left[\frac{M}{(M-1)P} + \Gamma\left(\frac{2M-1}{M-1}\right)\frac{2^{-\frac{B}{NK(M-1)}}}{P}\right.$$
$$\left.\times \left(\prod_{p=1}^N \prod_{q=1}^K \left(\frac{d_{p,q}^{-\alpha}}{\sigma_{p,q}^2}\right)^{\frac{1}{NK}}\right)C^\star\right]\sum_{i=1,i\neq n}^N \left(\frac{\sigma_{n,1}^2}{d_{n,1}^{-\alpha}}-\frac{\sigma_{i,1}^2}{d_{i,1}^{-\alpha}}\right), \tag{21}$$

---

[6]Specifically, certain level of ESR of each user needs to be achieved by power allocation in this paper.

[7]$\Delta R_{n,k}^{LB2}$ can be obtained exactly in closed-form. However, since it is not considered further for power allocation problem, we will not present the result.

where $C^\star$ is the solution of the following equation of variable C given as

$$C = P^{\frac{1}{K}} \left(\prod_{p=1}^{N} \left(1 - \phi_p^\star\right)^{\frac{1}{N}} \right)$$
$$\times \prod_{p=1}^{N} \prod_{q=2}^{K} \left(\frac{1}{P} + \frac{d_{p,q}^{-\alpha}(q-1)\phi_p^\star}{\sigma_{p,q}^2 K} \right)^{-\frac{1}{NK}} \quad (22)$$

with $\phi_p^\star$ $(p = 1, 2, \cdots, N)$ being a function of C given by (21). The desired solution of $C^\star$ can be obtained by solving the equation of C in (22) with any numerical method. Moreover, the power allocation satisfies $\phi_p^\star \geq \phi_2^\star \geq \cdots \geq \phi_N^\star$.

Proof: See Appendix G. □

Remark 5: Generally, there are multiple roots with the equation (22). The desired solution of $C^\star$ is chosen as the one makes the resulting $\phi_n^\star \in [0, 1]$ practical. If there are more than one roots of $C^\star$, the desired one can be determined by checking the obtained $\tilde{R}_{sum}^{LB2}$ (1).

According to the power allocation in Theorem 4, it is easy to see, given N the total transmit power P needs to be large enough to support N clusters of NOMA users based on our framework described above. Otherwise, $\phi_N^\star$ will become a negative number. When P is not large enough to support N clusters of users, there are two ways to circumvent this problem: (a) One can reduce the number of clusters by one and redo the user clustering of the NK users according to (2) and (3) until $\phi_N^\star > 0$ holds for the number of clusters N; or (b) remove the users in cluster N from the time-frequency resource block (RB) considered and multiplex these users on some other time-frequency RB.

C. Asymptotic Results of the Proposed Power and Feedback Bits Allocation

In order to provide insights for the system design, we now pursue an asymptotic analysis on our joint optimization of power and feedback bits allocation that are obtained above. First, the power allocation in the high power region and in the high CSI quantization accuracy scenario are given respectively as follows.

Corollary 1: When $B \to +\infty$ with fixed and finite P, our solution in *Theorem* 4 tends to be water-filling type solution for the system with perfect CSIT. Moreover, when $P \to +\infty$ with fixed and finite B, our power allocation solution tends to be equal power allocation among all users.

Proof: First, it can be observed from (21) that, when $B \to +\infty$ with fixed and finite P, our solution turns to be $\phi_n^\star \to \frac{1}{N} - \frac{K}{NP} \frac{M}{(M-1)} \sum_{i=1, i \neq n}^{N} \left(\frac{\sigma_{n,1}^2}{d_{n,1}^{-\alpha}} - \frac{\sigma_{i,1}^2}{d_{i,1}^{-\alpha}} \right)$, which is exactly the water-filling type solution for the system with perfect CSIT. Moreover, when $P \to +\infty$ with finite B,

$$C^\star \to \frac{P^{\frac{1}{K}} \left(\prod_{p=1}^{N} \left(1 - \phi_p^\star\right)^{\frac{1}{N}} \right)}{\left[\left(\prod_{p=1}^{N} \phi_p^\star \right)^{1-\frac{1}{K}} \prod_{p=1}^{N} \prod_{q=1}^{K} \left(\frac{(q-1)d_{p,q}^{-\alpha}}{\sigma_{p,q}^2 K} \right)^{\frac{1}{NK}} \right]}. \quad (23)$$

Then, the equal power allocation follows by substituting (23) into (21). □

Theorem 5: As the total transmit power P increases from the medium power region to high power region, with finite total bandwidth of feedback channels B and without integer constraint, the optimal number of feedback bits allocated to the nearest user $(n, 1)$ in each cluster is scaled approximately *increasingly* at rate $(M - 1)\left(1 - \frac{1}{K}\right) \log_2(P)$, whilst the optimal number of feedback bits allocated to the other users is scaled approximately *decreasingly* at rate $(M - 1)\left(1 - \frac{1}{K}\right) \log_2(P)$. Specifically, as P increases we have

$$B_{n,1}^\star \approx \tilde{B}_{n,1}^\star = \frac{B}{NK} + (M - 1)\left(1 - \frac{1}{K}\right) \log_2\left(\frac{P}{NK}\right)$$
$$+ (M - 1) \log_2\left(\frac{d_{n,1}^{-\alpha}}{\sigma_{n,1}^2}\right) - \frac{M - 1}{NK} \sum_{i=1}^{N} \log_2\left(\frac{d_{i,1}^{-\alpha}}{\sigma_{i,1}^2}\right)$$
$$+ \frac{M - 1}{K} \sum_{l=1}^{K-1} \log_2(l), \quad \forall n; \quad (24)$$

$$B_{n,k}^\star \approx \tilde{B}_{n,k}^\star = \frac{B}{NK} - \frac{M - 1}{K} \log_2\left(\frac{P}{NK}\right)$$
$$- \frac{M - 1}{NK} \sum_{i=1}^{N} \log_2\left(\frac{d_{i,1}^{-\alpha}}{\sigma_{i,1}^2}\right) + \frac{M - 1}{K} \sum_{l=1}^{K-1} \log_2(l)$$
$$- (M - 1) \log_2(k - 1), \quad \text{for } k \geq 2 \text{ and } \forall n. \quad (25)$$

Proof: According to *Corollary* 1, the power allocation of our joint optimization method tends to be equal power allocation in the high power region. Then, the theorem can be proved easily by substituting the result of $P_n = P/N$ into $\bar{S}_{n,k}^{(2)}$, $\bar{S}_{n,k}^{(3)}$ and (14) with some manipulations. □

Remark 6: We see from this theorem that the feedback bits allocated to each user by our method varies differently as P increases, whereas the feedback bits allocation by the method in [21] remains invariant as P varies. Moreover, this theorem shows, as P increases from the medium power region to high power region, the nearest (strongest) user in each cluster is allocated more feedback bits than the other users in the cluster, and among these nearest users of all clusters, the stronger the user is the more feedback bits is allocated. In contrast, for the users other than the strongest user in each cluster (i.e., user (n, k) for $k \geq 2$), the weaker the user is the more feedback bits is allocated. These insights can also be verified by the numerical results in Section V.

Moreover, we have the following result for finite B.

Corollary 2: With finite B and without integer constraint, as P goes to $+\infty$, the optimal number of feedback bits allocated to the users other than the nearest user in each cluster reduces to zero, whilst the optimal number of feedback bits allocated to user $(n, 1)$ of each cluster converges to a finite number. Specifically, we have $B_{n,1}^\star \approx \min\{\hat{B}_{n,1}^\star, \tilde{B}_{n,1}^\star\}$ and $B_{n,k}^\star \approx \min\{0, \tilde{B}_{n,k}^\star\}$ for $k \geq 2$, where each $\tilde{B}_{n,k}^\star$ is defined by (24) or (25) in *Theorem* 5 and

$$\hat{B}_{n,1}^\star \quad (26)$$
$$= \frac{B}{N} + (M - 1) \log_2\left(\frac{d_{n,1}^{-\alpha}}{\sigma_{n,1}^2}\right) - \frac{M - 1}{N} \sum_{i=1}^{N} \log_2\left(\frac{d_{i,1}^{-\alpha}}{\sigma_{i,1}^2}\right).$$

Proof: As $P \to +\infty$ with finite B, it is obvious that $B_{n,k}^\star \to 0$ and $\tilde{B}_{n,k}^\star \to 0$ for $k \geq 2$ and $\forall n$. The result of $\hat{B}_{n,1}^\star$ in (26) is obtained by substituting the expression of $\tilde{B}_{i,1}^\star$ in (24) into the constraint $\sum_{i=1}^N \tilde{B}_{i,1}^\star = B$, which completes the proof of the corollary. $\qquad\square$

Remark 7: It is easy to check that $\{\tilde{B}_{n,k}^\star\}$ satisfies the constraint on the total bandwidth of feedback channels. According to (24) and the user clustering criteria in (2) and (3), it is easy to see when P is large enough, our bits allocation results in $B_{1,1}^\star \geq B_{2,1}^\star \geq \cdots, \geq B_{N,1}^\star$.

In the following, we will show for our joint power and feedback bits allocation method, the proposed user clustering (denoted as $\mathfrak{C}$) in general is the optimal in terms of ESR among all possible user clustering methods. First note that, it is obvious that, if a different user clustering (denoted as $\tilde{\mathfrak{C}}$) is obtained from $\mathfrak{C}$ by exchanging users (i,p) and (j,q) for $i \neq j$ and $p \neq q$ where (i,p) and (j,q) follow (2) and (3), the ESR of $\tilde{\mathfrak{C}}$ will be reduced compared to that of $\mathfrak{C}$. Therefore, we can only focus on the user clustering obtained from $\mathfrak{C}$ by fixing the strongest user in each cluster and exchanging a few pairs of users with indices (i,k) and (j,k) for $i \neq j$ and $k \geq 2$. Then, we first consider $\tilde{\mathfrak{C}}$ obtained from $\mathfrak{C}$ by exchanging any two users (n,k) and (m,k) with $m \leq n$ (i.e., user (n,k) and user (m,k) in $\mathfrak{C}$ become user (m,k) and user (n,k) respectively in $\tilde{\mathfrak{C}}$). According to Theorem 3 and 4, it is easy to see the power allocated to user (n,k) and (m,k) originally in $\mathfrak{C}$ are exchanged and the feedback bits allocated to these two users are varied. However, the power and bits allocated to the other users do not change. Thus, the ergodic rate of any user other than the above two in $\tilde{\mathfrak{C}}$ is the same as that of the corresponding user in $\mathfrak{C}$.

We denote the ergodic rates of users (m,k) and (n,k) in $\tilde{\mathfrak{C}}$ as $\tilde{R}_{m,k}$ and $\tilde{R}_{n,k}$ respectively, which are obtained by substituting the power and feedback bits allocation results (without integer constraint) given by Theorem 3 and 4 into the expression of SINR in (5). Compared with $\mathfrak{C}$, since in $\tilde{\mathfrak{C}}$ the stronger user is allocated less power and the weaker user is allocated with more power (i.e., $P_n < P_m$), it is reasonable for one to anticipate the following result $R_{m,k} + R_{n,k} > \tilde{R}_{m,k} + \tilde{R}_{n,k}$ holds in general. In addition, the larger the gap of power allocation $P_m - P_n \geq 0$ is or equivalently the larger the distance between the two indices $n - m$ is, the larger the gap of ESR $(R_{m,k} + R_{n,k}) - (\tilde{R}_{m,k} + \tilde{R}_{n,k})$ becomes. Moreover, if $\tilde{\mathfrak{C}}$ is obtained from $\mathfrak{C}$ by exchanging more such user pairs, the rate gap between two user clusterings $\mathfrak{C}$ and $\tilde{\mathfrak{C}}$ will be increased. We find after some tedious manipulations that, the *exact* closed-form results of $R_{i,k}$ and $\tilde{R}_{i,k}$ for $i = m, n$ are necessary for the comparison between $\mathfrak{C}$ and $\tilde{\mathfrak{C}}$ *in theory*. But unfortunately, as we have noted in Section III that it is very difficult if not impossible to obtain the exact closed-form results. Thus, the ESR comparison between two user clusterings in theory is intractable. We have to turn to numerical experiments.

For verification, we compare the ESRs of different user clustering methods with the system parameters given by $M = 6$, $B = 72$ bits, $N = 3$ and $K = 2$. The AWGN at all users are with the equal power, i.e., $\sigma_{n,k}^2 = -50$ dBm $\forall n, k$. The distance parameters are given by $\mathbf{D}_1 =$
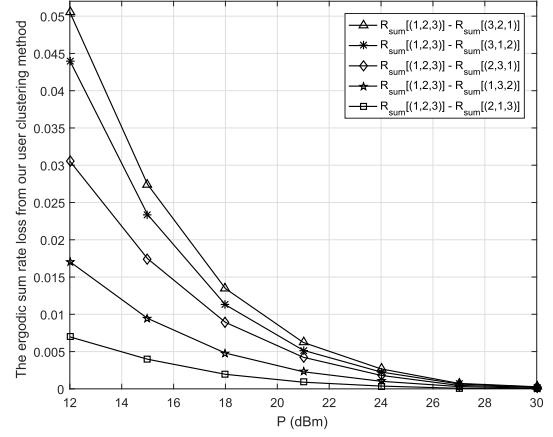

Fig. 1. The ESR losses of some of the other user clustering methods compared with our user clustering method.

$\begin{bmatrix} 25 & 35 \\ 27 & 37 \\ 29 & 39 \end{bmatrix}$ with the (n,k)-th element denoting $d_{n,k}$ in meters. Fig. 1 shows the ESR gaps between our user clustering the other methods, where $R_{\text{sum}}[(p,q,r)]$ with $p, q, r \in \{1,2,3\}$ denotes the ESR of $\tilde{\mathfrak{C}}$, and the users $(1,2), (2,2), (3,2)$ in $\tilde{\mathfrak{C}}$ corresponds to the users $(p,2), (q,2), (r,2)$ originally in $\mathfrak{C}$ respectively. Specifically, $R_{\text{sum}}[(1,2,3)]$ denotes ESR of our method. As anticipated, it is easy to see from Fig. 1 that the performance of our method is the optimal. Moreover, we can observe the following order of the ESR that $\{R_{\text{sum}}[(1,2,3)]\} > \{R_{\text{sum}}[(2,1,3)], R_{\text{sum}}[(1,3,2)]\} > \{R_{\text{sum}}[(3,1,2)], R_{\text{sum}}[(2,3,1)]\} > \{R_{\text{sum}}[(3,2,1)]\}$, which complies with the intuitive analysis above. However, the relative amounts of ESRs within a ESR set depend on the actual system parameters in general, and thus are *not comparable*. In the following section, we will validate the obtained analytical results and our proposed scheme by more numerical results.

V. NUMERICAL RESULTS

For the numerical results, we consider the systems with two different sets of distance parameters given as $\mathbf{D}_1$ of Fig. 1 and $\mathbf{D}_2 = \begin{bmatrix} 10 & 35 \\ 12 & 37 \\ 14 & 39 \end{bmatrix}$ and the different number of Bs. The other major system parameters are the same as those of Fig. 1. Specifically, the systems associated with Fig. 2 - Fig. 4 and the systems associated with Fig. 5 - Fig. 6 are with $\mathbf{D}_1$ and $\mathbf{D}_2 = \begin{bmatrix} 10 & 35 \\ 12 & 37 \\ 14 & 39 \end{bmatrix}$ respectively. The path-loss exponent $\alpha = 4$. Moreover, if the total transmit power P can not support a given number of clusters (i.e., $N = 3$ at the beginning of system design), we will follow the way of (a) in the design as we have described after Remark 5. The ESR curves corresponding to the obtained analytical lower bound in Theorem 1 will be notified specifically. Otherwise, the ESR curves are obtained from simulations.

Fig. 2 shows the simulation results of the ESR v.s P for our proposed joint bits and power allocation method and the corresponding analytical lower bound given by *Theorem* 1 with $B = 42$ bits. For comparisons, we also plot the simulation and analytical results with the quantized CSI model in [21]. We can see from the figure, the simulation results illustrate the fact we have explained in *Remark* 2, i.e., the channel model for limited feedback used in [21] actually underestimates the performance of the system with the well recognized practical
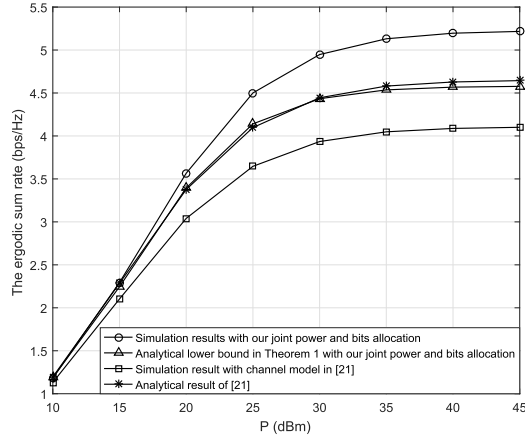
Fig. 2. Comparisons of the simulation and analytical results of the ESR v.s. P for our proposed scheme and those in [21].
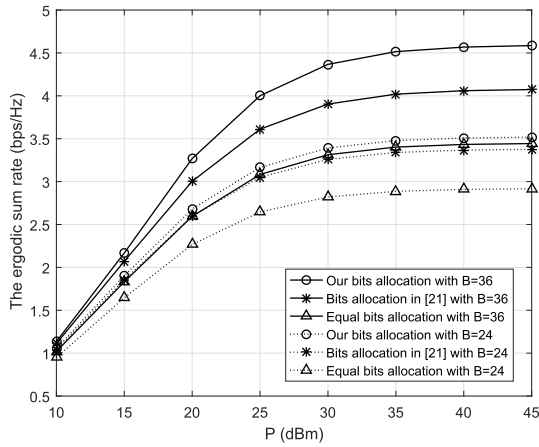

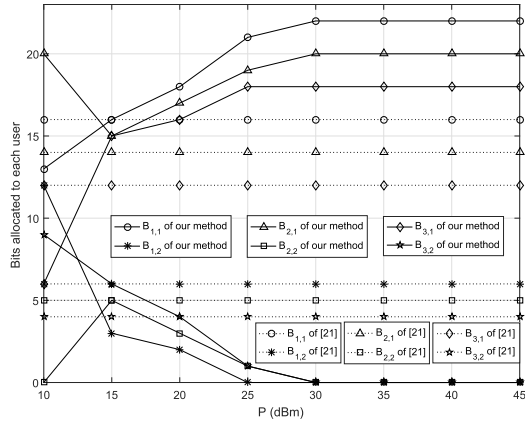Fig. 3. Comparisons of the ESR v.s P for different feedback bits allocation methods.

Fig. 4. Comparisons of the number of feedback bits allocated to each user for different bits allocation methods.

Fig. 5. Comparison of the system ESRs as a function of P with different power allocation methods.

Fig. 6. Comparisons of the ESRs as a function of B achieved by different power and feedback bits allocation methods.

impact of the aforementioned dependent relationships can not be ignored in the analysis and design of the practical systems. Moreover, as transmit power varies, our obtained analytical lower bound can track the real values well.

Fig. 3 compares the ESR as a function of P achieved by our proposed bits allocation method with those of the equal bits allocation and the bits allocation method proposed in [21] for the systems with different Bs. All schemes employ the quantized channel model in this paper and the equal power allocation among all users. We can see our bits allocation method performs the best of all. And the performance gaps between our method and the other two both increase when B becomes large, which illustrates the advantage of our method.

Fig. 4 shows the number of feedback bits allocated to each user obtained from our method and that obtained from the method in [21] with the total feedback channel bandwidth of $B = 60$ bits. We can see that, as P increases our method allocates more bits to the nearest users (to the BS) in each cluster, and allocates less bits to the farther users. In very high power region, our method does not allocate feedback bits for the farther user in each cluster. It is not difficult to check the results show in this figure are consistent with our analysis in Subsection IV-C. The results also can be

channel model used in many previous works [19], [20], [23]. However, we can see in the figure that the analytical ergodic rate result derived in [21] is right over the corresponding results of the simulations. This illustrates the analysis method in [21] to ignore the statistical dependence between the useful signal and MU interference signals received at each user results in *overestimation* of the actual performance. The big gap between analytical and simulated results demonstrates the
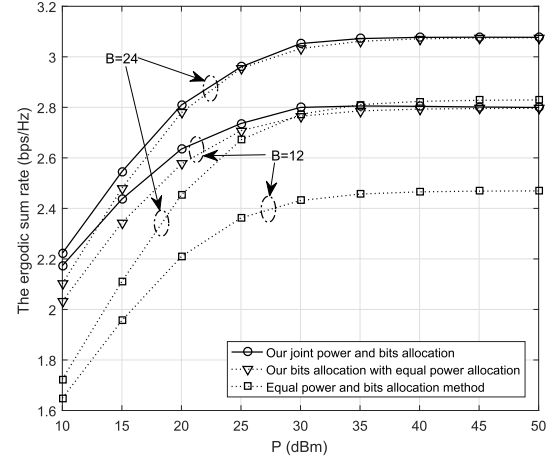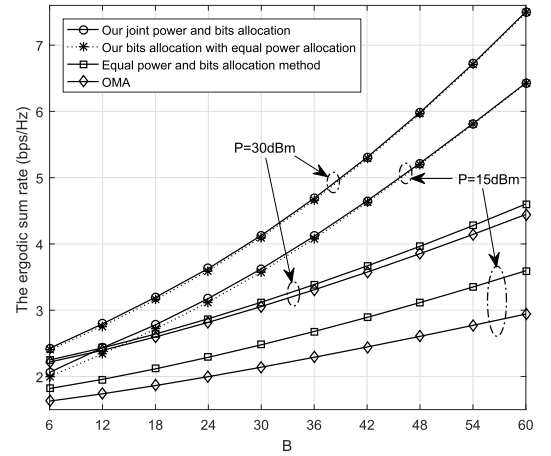
explained by the qualitative analysis as follows. On the one hand, increasing the quantization accuracy can enhance the strength of each user's transmit signal, and the increasing of the near users's signal strength is more effective in enhancing the transmission rate than increasing the farther users' signal strength. On the other hand, the system is interference-limited in high power region. Increasing the quantization accuracy of user (n, k)'s CSI can reduce the inter-cluster interference to all users (m, k) $\forall m \neq n$ and $\forall k$, but only reduces the intra-cluster interference to users (n, j) for $j > k$ due to SIC. Thus, with equal power allocation within each cluster, increasing the quantization accuracy of user (n, k)'s CSI is more effective in reducing the MU interference than increasing the quantization accuracy of user (n, j)'s CSI for $j > k$. In contrast, the bits allocation by the method of [21] does not change with the increase of P.

Fig. 5 shows the ESR as a function of P for different total bandwidths of feedback channels achieved respectively by our proposed joint optimization method, our bits allocation method with equal power allocation among all users, and the equal power and bits allocation method. As can be seen from the figure, our proposed power allocation converges to the equal power allocation as P goes large, which is consistent with our analysis in *Corollary* 1. However, there is always a gap between the performance of equal power and bits allocation and those of the other two.

Fig. 6 compares the ESR as a function of B achieved by our joint optimization method, our bits allocation method with equal power allocation and equal power and bits allocation method. As the baseline, we also plot the performance of the OMA transmission, where, for fair comparisons, the groups of the nearer users and the farther users share a fraction of $1/2$ time resource using linear ZFBF with equal power allocation for transmission. As can be seen from the figure, the performance of our joint bits and power allocation method is the best of all. The performances of all NOMA transmission schemes are better than that of OMA scheme for the system setting considered. Moreover, we can observe that, the performance gap between our joint optimization scheme and the equal bits allocation scheme and also the performance gap between our joint optimization scheme and the OMA scheme both increase as the accuracy of channel quantization (B) increases. However, the advantage of our power allocation method over the equal power allocation method decreases as the accuracy of channel quantization increases.

VI. CONCLUSION

In this paper, we have investigated the transmit-receive design, performance analysis, and optimization of downlink multi-antenna NOMA cellular networks with general limited feedback in FDD mode. A mathematically strict performance analysis of the ergodic rate for the considered system was conducted for the first time. A closed-form lower bound on the ergodic rate of each user was obtained in terms of transmit power, CSI quantization accuracy and channel conditions. Then, we optimized the two key parameters of the considered system, i.e., transmit power and the number of feedback bits allocated to each users. Finally, we conducted asymptotic performance analysis on the optimization results, and obtained insights on system design guidelines.

APPENDIX

A. Proof of Theorem 1

According to [23, *Lemma* 2], we have $|\tilde{\mathbf{e}}_{n,k}\mathbf{w}_i|^2 \sim$ Beta$(1, M - 2)$ and $\mathbb{E}\left[|\tilde{\mathbf{e}}_{n,k}\mathbf{w}_i|^2\right] = (M - 1)^{-1}$. Then, by employing Jensen's inequality we can lower-bound $R_{n,k}$ for $k \geq 2$ as (29) at the bottom of this page, where $Z \triangleq |\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n|^2 \sim$ Beta$(1, M - 1)$ [23]. It has been shown in [24, *Lemma* 2] that, with cell approximation of vector quantization the joint distribution of the random variable (RV) pair $(||\mathbf{h}_{n,k}||^2 \sin^2 \theta_{n,k}, ||\mathbf{h}_{n,k}||^2 \cos^2 \theta_{n,k})$ is the same as that of (I, S), where $I = \delta_{n,k}Y$ and $S = X + (1 - \delta_{n,k})Y$ with $X \sim$ Gamma$(1, 1)$ and $Y \sim$ Gamma$(M - 1, 1)$ being two independent RVs. Since X, Y and Z are independent of each other [23], $R_{n,k}^{LB1}$ in (29) can be obtained as

$$
\begin{aligned}
R_{n,k}^{LB1} = \mathbb{E}\Big\{ &\log_2 \Big[1 + S_{n,k}^{(1)} Z(X + (1 - \delta_{n,k}Y)) \\
&+ \Big(S_{n,k}^{(1)} Z + S_{n,k}^{(3)} (M - 1)^{-1}\Big)\delta_{n,k}Y\Big]\Big\} \\
- \mathbb{E}\Big\{ &\log_2 \Big[1 + S_{n,k}^{(2)} Z(X + (1 - \delta_{n,k}Y)) \\
&+ \Big(S_{n,k}^{(1)} Z + S_{n,k}^{(3)} (M - 1)^{-1}\Big)\delta_{n,k}Y\Big]\Big\} \\
= \mathbb{E}_Z \Big[&\Upsilon_{n,k}\Big(S_{n,k}^{(1)} Z, \ S_{n,k}^{(1)} Z + S_{n,k}^{(3)} (M - 1)^{-1}\delta_{n,k}\Big)\Big] \\
- \mathbb{E}_Z \Big[&\Upsilon_{n,k}\Big(S_{n,k}^{(2)} Z, \ S_{n,k}^{(2)} Z + S_{n,k}^{(3)} (M-1)^{-1}\delta_{n,k}\Big)\Big], \quad (30)
\end{aligned}
$$

where $\Upsilon_{n,k}(\mu, \nu) \triangleq \mathbb{E}_{X,Y}\left[\log_2(1 + \mu X + \nu Y)\right]$. The result of $\mathbb{E}_Z\left[\Upsilon_{n,k}(aZ, aZ + b)\right] \triangleq \Theta_{n,k}(a, b)$ is derived in *Appendix* B, where $\Theta_{n,k}(a, b)$ is given in (8). Then, the result of $R_{n,k}^{LB1}$ in (7) follows.

$$
R_{n,k} \geq \mathbb{E}\left[\log_2\left(1 + \frac{\frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2}|\mathbf{h}_{n,k}\mathbf{w}_n|^2 \ P_{n,k}}{1 + S_{n,k}^{(2)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2 + S_{n,k}^{(3)}\mathbb{E}\left[|\tilde{\mathbf{e}}_{n,k}\mathbf{w}_i|^2\right]||\mathbf{h}_{n,k}||^2 \sin^2 \theta_{n,k}}\right)\right] \tag{27}
$$

$$
= \mathbb{E}\left[\log_2\left(\frac{1 + S_{n,k}^{(1)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2 + S_{n,k}^{(3)}(M - 1)^{-1}||\mathbf{h}_{n,k}||^2 \sin^2 \theta_{n,k}}{1 + S_{n,k}^{(2)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2 + S_{n,k}^{(3)}(M - 1)^{-1}||\mathbf{h}_{n,k}||^2 \sin^2 \theta_{n,k}}\right)\right] \tag{28}
$$

$$
= \mathbb{E}\left[\log_2\left(1 + S_{n,k}^{(1)}Z||\mathbf{h}_{n,k}||^2 + S_{n,k}^{(3)}(M - 1)^{-1}||\mathbf{h}_{n,k}||^2 \sin^2 \theta_{n,k}\right)\right]
$$
$$
- \mathbb{E}\left[\log_2\left(1 + S_{n,k}^{(2)}Z||\mathbf{h}_{n,k}||^2 + S_{n,k}^{(3)}(M - 1)^{-1}||\mathbf{h}_{n,k}||^2 \sin^2 \theta_{n,k}\right)\right] \triangleq R_{n,k}^{LB1}, \tag{29}
$$

When $k = 1$, $S_{n,k}^{(2)} = 0$. By replacing $S_{n,k}^{(2)}$ in (30) with 0, we can similarly obtain

$$
\begin{aligned}
R_{n,1}^{LB1} &= \mathbb{E}\Big\{ \log_2 \Big[1 + \alpha_{n,1} P_{n,1} ZX \\
&\quad + \Big(\alpha_{n,1} P_{n,1} Z + S_{n,1}^{(3)} (M-1)^{-1} \delta_{n,1} \Big) Y \Big] \Big\} \\
&\quad - \mathbb{E}\Big\{ \log_2 \Big[1 + S_{n,1}^{(3)} (M-1)^{-1} \delta_{n,1} Y \Big] \Big\} \\
&= \Theta_{n,k}(\alpha_{n,1} P_{n,1}, S_{n,1}^{(3)} (M-1)^{-1} \delta_{n,1}) \\
&\quad - \mathbb{E}\Big[\log_2 \Big(1 + S_{n,1}^{(3)} (M-1)^{-1} \delta_{n,1} Y \Big) \Big]. \quad (31)
\end{aligned}
$$

The second term in (31) can be obtained as

$$
\begin{aligned}
&\mathbb{E}\Big[\log_2 \Big(1 + S_{n,1}^{(3)} (M-1)^{-1} \delta_{n,1} Y \Big) \Big] \\
&= \int_0^{+\infty} \log_2 \Big(1 + S_{n,1}^{(3)} (M-1)^{-1} \delta_{n,1} y \Big) \frac{e^{-y} y^{M-2}}{(M-2)!} \, dy \\
&= \log_2(e) \exp\Big(\frac{M-1}{S_{n,1}^{(3)} \delta_{n,1}} \Big) \sum_{q=1}^{M-1} E_q \Big(\frac{M-1}{S_{n,1}^{(3)} \delta_{n,1}} \Big), \quad (32)
\end{aligned}
$$

where (32) follows using (34) in *Appendix* B. Then, $R_{n,k}^{LB1}$ in (6) for $k = 1$ follows by substituting (32) into (31).

B. The Derivation of $\mathbb{E}_Z[\Upsilon_{n,k}(aZ, aZ + b)] \triangleq \Theta_{n,k}(a,b)$ for $a > 0$, $b > 0$

Using the result in [28], the probability density function (PDF) of $J_{\mu,\nu} = \mu X + \nu Y$ for $\mu \neq \nu$ is given by $f_{J_{\mu,\nu}}(x) = \frac{(-1)^{M-1} \mu^{M-2}}{(\nu-\mu)^{M-1}} e^{-\frac{x}{\mu}} + \sum_{p=1}^{M-1} \frac{(-\mu)^{p-1}}{(M-1-p)!(\nu-\mu)^p \nu^{M-1-p}} x^{M-1-p} e^{-\frac{x}{\nu}}$ $(x > 0)$. Thus,

$$
\begin{aligned}
\Upsilon_{n,k}(\mu,\nu) &= \mathbb{E}_{X,Y}[\log_2(1 + \mu X + \nu Y)] \\
&= \int_0^{+\infty} \log_2(1+x) f_{J_{\mu,\nu}}(x) \, dx \\
&= \log_2(e) \Big[\frac{(-\mu)^{M-1}}{(\nu-\mu)^{M-1}} \exp\Big(\frac{1}{\mu} \Big) E_1 \Big(\frac{1}{\mu} \Big) \\
&\quad + \sum_{p=1}^{M-1} \frac{(-\mu)^{p-1}\nu}{(\nu-\mu)^p} \exp\Big(\frac{1}{\nu} \Big) \sum_{q=1}^{M-p} E_q \Big(\frac{1}{\nu} \Big) \Big], \quad (33)
\end{aligned}
$$

where (33) follows from the result of [29, (78)]

$$
\begin{aligned}
&\int_0^{+\infty} \ln(1 + ax) e^{-\mu x} x^{p-1} dx \\
&\qquad\qquad = (p-1)! e^{\frac{\mu}{a}} \mu^{-p} \sum_{q=1}^{p} E_q \Big(\frac{\mu}{a} \Big). \quad (34)
\end{aligned}
$$

Then, substituting $\mu = aZ$ and $\nu = aZ + b$ with $a, b > 0$ into (33), we can obtain

$$
\begin{aligned}
&\mathbb{E}_Z[\Upsilon_{n,k}(aZ, aZ+b)] \\
&= \log_2(e) \mathbb{E}_Z \Big[\frac{(-aZ)^{M-1}}{b^{M-1}} \exp\Big(\frac{1}{aZ} \Big) E_1 \Big(\frac{1}{aZ} \Big) + \sum_{p=1}^{M-1} \\
&\frac{(-aZ)^{p-1}(aZ+b)}{b^p} \exp\Big(\frac{1}{aZ+b} \Big) \sum_{q=1}^{M-p} E_q \Big(\frac{1}{aZ+b} \Big) \Big]. (35)
\end{aligned}
$$

By using the relationship [26] $E_q(x) = \frac{1}{q-1}[e^{-x} - x E_{q-1}(x)]$ and $E_1(x) = -E_i(-x)$ with induction, $E_q(x)$ can be expressed with $E_i(-x)$ as $E_q(x) = e^{-x} \sum_{s=0}^{q-2} \frac{(q-s-2)!(-x)^s}{(q-1)!} + \frac{(-1)^q (x)^{q-1}}{(q-1)!} E_i(-x)$. Then, substituting this result into (35) with the PDF of Z, we can obtain

$$
\begin{aligned}
&\mathbb{E}_Z[\Upsilon_{n,k}(aZ, aZ+b)] \\
&= \log_2(e) \mathbb{E}_Z \Big[\frac{(-1)^M (aZ)^{M-1}}{b^{M-1}} \exp\Big(\frac{1}{aZ} \Big) E_i \Big(-\frac{1}{aZ} \Big) \\
&\quad + \sum_{p=1}^{M-1} \sum_{q=1}^{M-p} \frac{(-1)^{p+q+1}(aZ)^{p-1}}{(q-1)! b^p (aZ+b)^{q-2}} \\
&\quad \times \exp\Big(\frac{1}{aZ+b} \Big) E_i \Big(-\frac{1}{aZ+b} \Big) \\
&\quad + \sum_{p=1}^{M-1} \sum_{q=2}^{M-p} \sum_{s=0}^{q-2} \frac{(-1)^{p+s-1}(q-s-2)!(aZ)^{p-1}}{(q-1)! b^p (aZ+b)^{s-1}} \Big], \quad (36)
\end{aligned}
$$

which is given by (8) with

$$
I_1(a) \triangleq \int_0^1 (az)^{M-1}(1-z)^{M-2} \exp\Big(\frac{1}{az} \Big) E_i \Big(-\frac{1}{az} \Big) dz, \quad (37)
$$

$$
\begin{aligned}
I_2(a,b,p,q) &\triangleq \int_0^1 \frac{(az)^{p-1}(1-z)^{M-2}}{(az+b)^{q-2}} \exp\Big(\frac{1}{az+b} \Big) \\
&\quad \times E_i \Big(-\frac{1}{az+b} \Big) dz, \quad (38)
\end{aligned}
$$

$$
I_3(a,b,p,s) \triangleq \int_0^1 \frac{(az)^{p-1}(1-z)^{M-2}}{(az+b)^{s-1}} dz. \quad (39)
$$

Then, we derive the results of (37), (38) and (39) as follows. For the convenience of presentation, we first define an integral $\Psi(n,u,v) \triangleq \int_u^v x^n e^x E_i(-x) dx$, whose closed-form result is given by (9) for different cases. The derivation is provided in *Appendix* C. Then, $I_1(a)$ can be obtained as

$$
\begin{aligned}
I_1(a) &= \sum_{t=0}^{M-2} \binom{M-2}{t} \frac{(-1)^t}{a^t} \int_0^1 (az)^{M-1+t} \\
&\qquad\qquad\qquad \times \exp\Big(\frac{1}{az} \Big) E_i \Big(-\frac{1}{az} \Big) dz \\
&= \sum_{t=0}^{M-2} \binom{M-2}{t} \frac{(-1)^t}{a^t} \int_{+\infty}^{\frac{1}{a}} x^{-M+1-t} e^x E_i(-x) \, d\Big(\frac{1}{ax} \Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad (40) \\
&= \sum_{t=0}^{M-2} \binom{M-2}{t} \frac{(-1)^t}{a^{t+1}} \Psi\Big(-M-1-t, \frac{1}{a}, v \Big) \Big|_{v \to +\infty},
\end{aligned}
$$

where (40) follows by the change of variables $x = \frac{1}{az}$. Similarly, $I_2(a,b,p,q)$ can be obtained by using the change of variables $x = \frac{1}{az+b}$ as

$$
I_2(a,b,p,q) = \sum_{r=0}^{M-2} \sum_{t=0}^{p-1+r} \binom{M-2}{r} \binom{p-1+r}{t}
$$

$$\times \frac{(-1)^{p-1-t}b^{p-1+r-t}}{a^{r+1}} \int_{\frac{1}{a+b}}^{\frac{1}{b}} x^{q-4-t}e^x E_i(-x) \ dx$$

$$= \sum_{r=0}^{M-2}\sum_{t=0}^{p-1+r}\binom{M-2}{r}\binom{p-1+r}{t}\frac{(-1)^{p-1-t}b^{p-1+r-t}}{a^{r+1}}$$

$$\times \Psi\left(q-4-t,\frac{1}{a+b},\frac{1}{b}\right).$$

Finally, $I_3(a,b,p,s)$ can be obtained as

$$I_3(a,b,p,s)$$

$$= \sum_{t=0}^{M-2}\binom{M-2}{t}\frac{(-1)^t a^{-t-1}}{b^{s-1}}\int_0^a \frac{x^{p-1+t}}{(\frac{x}{b}+1)^{s-1}}dx \quad (41)$$

$$= \sum_{t=0}^{M-2}\binom{M-2}{t}\frac{(-1)^t a^{p-1}}{b^{s-1}(p+t)}$$

$$\times {}_2F_1\left(s-1,p+t;p+t+1;-\frac{a}{b}\right), \quad (42)$$

where (41) is obtained from the change of variables $x = az$ and (42) follows from [26, 3.194.1] $\int_0^u \frac{x^{\mu-1}dx}{(1+\beta x)^\nu} = \frac{u^\mu}{\mu}{}_2F_1(\nu,\mu;1+\mu;-\beta u)$. Then, the final results can be obtained by substituting the results of $I_1(a)$, $I_2(a,b,p,q)$ and $I_3(a,b,p,s)$ into (36).

C. Derivation of $\Psi(n,u,v)$

For $n > 0$, using the results [26, 2.321.2] $\int x^n e^{ax}dx = e^{ax}\left(\sum_{k=0}^n \frac{(-1)^k n!}{a^{k+1}(n-k)!}x^{n-k}\right)$ and $(E_i(-x))' = x^{-1}e^{-x}$, $\Psi(n,u,v)$ can be obtained by integration by parts as

$$\Psi(n,u,v) = \int_u^v E_i(-x) \ d\left[e^x\sum_{k=0}^n \frac{(-1)^k n!}{(n-k)!}x^{n-k}\right] \quad (43)$$

$$= \left[E_i(-x)e^x\sum_{k=0}^n \frac{(-1)^k n!}{(n-k)!}x^{n-k}\right]_u^v$$

$$- \sum_{k=0}^n \frac{(-1)^k n!}{(n-k)!}\int_u^v x^{n-k-1} \ dx, \quad (44)$$

which is given in (9) with $n > 0$.

For $n \leq -2$, using integration by part again with the result [26, 2.324.2] $\int \frac{e^{ax}}{x^n}dx = -e^{ax}\sum_{k=1}^{n-1}\frac{(n-k-1)!a^{k-1}}{(n-1)!x^{n-k}} + \frac{a^{n-1}E_i(ax)}{(n-1)!}(n>0)$ for $a=1$, $\Psi(n,u,v)$ can be obtained as

$$\Psi(n,u,v) = \left[E_i(-x)\left(-e^x\sum_{k=1}^{-n-1}\frac{(-n-k-1)!}{(-n-1)!}x^{n+k}\right.\right.$$

$$\left.\left.+\frac{E_i(x)}{(-n-1)!}\right)\right]_u^v$$

$$- \int_u^v\left(-e^x\sum_{k=1}^{-n-1}\frac{(-n-k-1)!}{(-n-1)!}x^{n+k}\right.$$

$$\left.+\frac{E_i(x)}{(-n-1)!}\right) d[E_i(-x)], \quad (45)$$

Moreover, $\int_u^v E_i(x) \ d[E_i(-x)]$ can be obtained as

$$\int_u^v E_i(x) \ d[E_i(-x)] = \int_u^v \left(E_i(-x)\right.$$

$$\left.+\sum_{m=0}^\infty \frac{2x^{2m+1}}{(2m+1)(2m+1)!}\right) d[E_i(-x)] \quad (46)$$

$$= \frac{1}{2}\left[E_i^2(-v)-E_i^2(-u)\right] + \sum_{m=0}^\infty\sum_{l=0}^{2m}\frac{2(u^l e^{-u}-v^l e^{-v})}{(2m+1)^2\,l!}, \quad (47)$$

where (46) follows from [26, 8.214.3] $E_i(x) = E_i(-x) + \sum_{m=0}^\infty \frac{2\,x^{2m+1}}{(2m+1)(2m+1)!}$ and (47) follows from [26, 3.351.1]. Then, the expression of $\Psi(n,u,v)$ for $n \leq -2$ in (9) can be obtained by combing (45) and (47). The results for $n = 0$ and $n = -1$ in (9) can be similarly obtained.

For $n < -1$ with $v \to +\infty$, we can obtain

$$\tilde{\Psi}(n,u)$$

$$= \left.\Psi(n,u,v)\right|_{v\to+\infty} = \int_u^\infty x^n e^x E_i(-x) \ dx$$

$$= \left[E_i(-x)\left(-e^x\sum_{k=1}^{-n-1}\frac{(-n-k-1)!}{(-n-1)!}x^{n+k}\right.\right.$$

$$\left.\left.+\frac{E_i(x)}{(-n-1)!}\right)\right]_u^\infty + \left[\sum_{k=1}^{-n-1}\frac{(-n-k-1)!\,x^{n+k}}{(-n-1)!(n+k)}\right]_u^\infty$$

$$- \frac{1}{(-n-1)!}\int_u^\infty E_i(x) \ d[E_i(-x)]. \quad (48)$$

Using [30, 5.1.19] and L'Hospital rule, after some manipulations it is easy to obtain $\lim_{x\to+\infty}e^x E_i(-x) = 0$ and $\lim_{x\to+\infty}E_i(-x)E_i(x) = 0$. Then, the first term in (48) can be obtained as $\left[E_i(-x)\left(-e^x\sum_{k=1}^{-n-1}\frac{(-n-k-1)!}{(-n-1)!}x^{n+k}+\frac{E_i(x)}{(-n-1)!}\right)\right]_u^\infty = E_i(-u)\left(e^u\sum_{k=1}^{-n-1}\frac{(-n-k-1)!}{(-n-1)!}u^{n+k}-\frac{E_i(u)}{(-n-1)!}\right)$. The second term in (48) is given by $\left[\sum_{k=1}^{-n-1}\frac{(-n-k-1)!\,x^{n+k}}{(-n-1)!(n+k)}\right]_u^\infty = -\sum_{k=1}^{-n-1}\frac{(-n-k-1)!\,u^{n+k}}{(-n-1)!(n+k)}$. Using [26, 3.351.2], the third term in (48) can be obtained as $\int_u^{+\infty}E_i(x)\ d[E_i(-x)] = -\frac{1}{2}E_i^2(-u)+\sum_{m=0}^\infty\sum_{l=0}^{2m}\frac{2\,u^l e^{-u}}{(2m+1)^2\,l!}$. Then, the result of $\left.\Psi(n,u,v)\right|_{v\to+\infty}$ in (10) follows by substituting the above results into (48).

D. Proof of Theorem 2

With the expressions of $R_{n,k}^{ideal}$ and $R_{n,k}$ we have

$$\Delta R_{n,k} = \mathbb{E}\left[\log_2\left(1+S_{n,k}^{(1)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2\right)\right]$$

$$-\mathbb{E}\left[\log_2\left(1+S_{n,k}^{(2)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2\right)\right]-\mathbb{E}\left[\log_2\left(1+S_{n,k}^{(1)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2\right.\right.$$

$$\left.\left.+\frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2}\sin^2\theta_{n,k}||\mathbf{h}_{n,k}||^2\sum_{i=1,i\neq n}^N |\tilde{\mathbf{e}}_{n,k}\mathbf{w}_i|^2\sum_{l=1}^K P_{i,l}\right)\right]$$

$$+\mathbb{E}\left[\log_2\left(1+S_{n,k}^{(2)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2+\frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2}\sin^2\theta_{n,k}||\mathbf{h}_{n,k}||^2\right.\right.$$

$$\left.\left.\times\sum_{i=1,i\neq n}^N |\tilde{\mathbf{e}}_{n,k}\mathbf{w}_i|^2\sum_{l=1}^K P_{i,l}\right)\right] \quad (49)$$

$$\leq \mathbb{E}\left[\log_2\left(1+S_{n,k}^{(2)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2+\frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2}\sin^2\theta_{n,k}||\mathbf{h}_{n,k}||^2\right.\right. \quad (50)$$

$$\left.\left.\times\sum_{i=1,i\neq n}^{N}|\tilde{\mathbf{e}}_{n,k}\mathbf{w}_i|^2\sum_{l=1}^{K}P_{i,l}\right)\right]-\mathbb{E}\left[\log_2\left(1+S_{n,k}^{(2)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2\right)\right]$$

$$\leq \log_2\left(1+S_{n,k}^{(2)}+\frac{M}{M-1}2^{B_{n,k}}\beta\left(2^{B_{n,k}},\frac{M}{M-1}\right)S_{n,k}^{(3)}\right)$$

$$-\log_2(e)e^{\frac{1}{S_{n,k}^{(2)}}}E_1\left(\frac{1}{S_{n,k}^{(2)}}\right)$$

$$=\log_2\left(1+S_{n,k}^{(2)}+\Gamma\left(\frac{2M-1}{M-1}\right)2^{-\frac{B_{n,k}}{M-1}}S_{n,k}^{(3)}\right)$$

$$-\log_2(e)e^{\frac{1}{S_{n,k}^{(2)}}}E_1\left(\frac{1}{S_{n,k}^{(2)}}\right), \quad (51)$$

where (50) follows by neglecting $\frac{d_{n,k}^{-\alpha}}{\sigma_{n,k}^2}\sin^2\theta_{n,k}||\mathbf{h}_{n,k}||^2$ $\times\sum_{i=1,i\neq n}^{N}|\tilde{\mathbf{e}}_{n,k}\mathbf{w}_i|^2\sum_{l=1}^{K}P_{i,l}$ in the third $\log_2$ of (49). (50) is obtained by applying Jensens inequality to the first $\log_2$ term of (50). (51) follows by combing the results $\mathbb{E}[|\mathbf{h}_{n,k}\mathbf{w}_n|^2]=1$, $\mathbb{E}[||\mathbf{h}_{n,k}||^2]=M$ and $\mathbb{E}[\sin^2\theta_{n,k}]=2^{B_{n,k}}\beta\left(2^{B_{n,k}},\frac{M}{M-1}\right)$ with $2^{B_{n,k}}\beta\left(2^{B_{n,k}},\frac{M}{M-1}\right)\approx\Gamma\left(\frac{M}{M-1}\right)2^{-\frac{B_{n,k}}{M-1}}$ and $x\Gamma(x)=\Gamma(x+1)$, which were shown in [23].

E. Proof of Theorem 3

It is easy to show that the objective function is logarithm convex function of variables $B_{n,k}$s. Thus, the problem of (13) is a convex optimization problem. The Lagrangian is given by

$$\mathcal{L}(B_{n,k},\lambda)=\lambda\left(\sum_{n=1}^{N}\sum_{k=1}^{K}B_{n,k}-B\right)$$

$$+\prod_{n=1}^{N}\prod_{k=1}^{K}\left(1+S_{n,k}^{(2)}+\Gamma\left(\frac{2M-1}{M-1}\right)2^{-\frac{B_{n,k}}{M-1}}S_{n,k}^{(3)}\right),$$

where λ is the is the Lagrange multiplier associated with total feedback bits constraint. Applying the Karush-Kuhn-Tucker (KKT) condition, we can get the necessary and sufficient conditions as

$$\left.\frac{\partial\mathcal{L}}{\partial B_{n,k}}\right|_{B_{n,k}^{\star},\,\lambda^{\star}}=-\frac{\ln(2)\Gamma\left(\frac{2M-1}{M-1}\right)S_{n,k}^{(3)}2^{-\frac{B_{n,k}}{M-1}}}{(M-1)}$$

$$\times\prod_{i=1}^{N}\prod_{j=1}^{K}\left(1+S_{i,j}^{(2)}+\Gamma\left(\frac{2M-1}{M-1}\right)2^{-\frac{B_{i,j}^{\star}}{M-1}}S_{i,j}^{(3)}\right)$$

$$\times\left(1+S_{n,k}^{(2)}+\Gamma\left(\frac{2M-1}{M-1}\right)2^{-\frac{B_{n,k}^{\star}}{M-1}}S_{n,k}^{(3)}\right)^{-1}+\lambda^{\star}=0,$$

$$\forall\,n,k, \quad (52)$$

$$\left.\frac{\partial\mathcal{L}(B_{n,k},\lambda)}{\partial\lambda}\right|_{B_{n,k}^{\star},\,\lambda^{\star}}=\sum_{n=1}^{N}\sum_{k=1}^{K}B_{n,k}^{\star}-B=0,$$

$$\text{if}\quad\lambda^{\star}\geq0. \quad (53)$$

From (52) we can obtain the relationship $B_{p,q}^{\star}=B_{n,k}^{\star}-(M-1)\log_2\left[\frac{S_{n,k}^{(3)}(1+S_{p,q}^{(2)})}{S_{p,q}^{(3)}(1+S_{n,k}^{(2)})}\right]$ for $\forall p\neq n$ and $\forall q\neq k$. Then, the solution of $B_{n,k}^{\star}$ in (14) follows by substituting this relationship into (53).

F. Derivation of the Lower Bound $R_{n,k}^{LB2}$

Based on (28), we have (54) and (55) at the top of the next page, where (54) follows from applying Jensens inequality to $|\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n|^2$ and the result $\mathbb{E}[|\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n|^2]=M^{-1}$ [23]. Then, by applying $\mathbb{E}[\log_2(1+X)]\geq\log_2(1+\frac{1}{\mathbb{E}[1/X]})$ to (55), we have $R_{n,k}$, shown at the top of the next page, which is given by (17) and the final result follows from the approximation of $\mathbb{E}[\sin^2\theta_{n,k}]$ in [23]. For $\Delta R_{n,k}^{LB2}$ in (17), we have

$$\Delta R_{n,k}^{LB2}-\mathbb{E}\left[\log_2\left(M|\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n|^2\right)\right]$$

$$=\mathbb{E}\left[\log_2\left(\frac{1+\bar{S}_{n,k}^{(1)}||\mathbf{h}_{n,k}||^2|\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n|^2}{\left(M+\bar{S}_{n,k}^{(1)}||\mathbf{h}_{n,k}||^2\right)|\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n|^2}\right)\right]$$

$$\geq\mathbb{E}\left[\log_2\left(\frac{1+\bar{S}_{n,k}^{(1)}||\mathbf{h}_{n,k}||^2\,M^{-1}}{\left(M+\bar{S}_{n,k}^{(1)}||\mathbf{h}_{n,k}||^2\right)M^{-1}}\right)\right]=0,$$

which follows from Jensen's inequality, since $\log_2\left(\frac{1+\zeta x}{(M+\zeta)x}\right)$ is a convex function of x.

Define $g(P_n)\triangleq\mathbb{E}\left[\log_2\left(1+\bar{S}_{n,k}^{(1)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2\right)\right]-\mathbb{E}\left[\log_2\left(1+\bar{S}_{n,k}^{(1)}M^{-1}||\mathbf{h}_{n,k}||^2\right)\right]$. It is easy to obtain $g'(P_n)\triangleq\frac{dg(P_n)}{dP_n}=\mathbb{E}\left[\left(1+\frac{d_{n,k}^{-\alpha}kP_n}{\sigma_{n,k}^2\,KM}||\mathbf{h}_{n,k}||^2\right)^{-1}\right]-\mathbb{E}\left[\left(1+\frac{d_{n,k}^{-\alpha}kP_n}{\sigma_{n,k}^2\,K}||\mathbf{h}_{n,k}||^2|\mathbf{h}_{n,k}\mathbf{w}_n|^2\right)^{-1}\right]$. Applying Jensens inequality on the second term, it is easy to show $g'(P_n)\leq0$. Thus, $g(P_n)$ is a decreasing function of P_n with $g(0)=0$. Thus, $g(P_n)<0$ when $P_n>0$. Moreover, since $g'(0)=0$ and $\lim_{P_n\to\infty}g'(P_n)=0$, $|g(P_n)|$ is relatively small when P_n is small and becomes less sensitive to P_n as P_n goes large. In addition, since $|\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n|^2\sim\text{Beta}(1,M-1)$, we have $\mathbb{E}\left[\log_2\left(M|\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n|^2\right)\right]=\log_2(M)+(M-1)\int_0^1\log_2(x)(1-x)^{M-2}dx=\log_2(M)-\log_2(e)\sum_{k=1}^{M-1}k^{-1}$ [26, 4.253.1, 8.365.3] . It is easy to show $h(M)=\log_2(M)-\log_2(e)\sum_{k=1}^{M-1}k^{-1}$ is an increasing function of M and $\lim_{M\to\infty}h(M)=-\log_2(e)\mathbf{C}\approx-0.8327$ [26, 8.367.2].

G. Proof of Theorem 4

First, denote $C\triangleq P^{\frac{1}{K}}\left(\prod_{p=1}^{N}(1-\phi_p)^{\frac{1}{N}}\right)\times\left(\prod_{p=1}^{N}\prod_{q=2}^{K}\left(\frac{1}{P}+\frac{d_{p,q}^{-\alpha}(q-1)\phi_p}{\sigma_{p,q}^2\,K}\right)^{-\frac{1}{NK}}\right)$. Then, by substituting C into (19) we can re-express $\tilde{R}_{n,k}^{LB2}(\{\phi_n\})\,\forall n,k$ as (56) shown at the top of the next page.

In the following, we identify C as a constant. Then, the problem in (20) becomes a convex optimization problem. The Lagrangian associated with the problem in (20) is defined as (57) shown at the top of the next page, $\forall n,k$ at the top of the next page, where λ_ϕ is the Lagrangian multiplier associated with the sum power constraint. According to the KKT conditions for optimality, we can obtain (58) and (59) at the top of the next page .

$$R_{n,k} \geq \mathbb{E}\left[\log_2\left(1 + \bar{S}_{n,k}^{(1)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2\right)\right] + \mathbb{E}\left[\log_2\left(1 + \frac{\bar{S}_{n,k}^{(3)}(M-1)^{-1}||\mathbf{h}_{n,k}||^2\sin^2\theta_{n,k}}{1 + \bar{S}_{n,k}^{(1)}\left|\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n\right|^2||\mathbf{h}_{n,k}||^2}\right)\right]$$

$$-\mathbb{E}\left[\log_2\left(1 + \bar{S}_{n,k}^{(2)}\left|\tilde{\mathbf{h}}_{n,k}\mathbf{w}_n\right|^2||\mathbf{h}_{n,k}||^2 + \bar{S}_{n,k}^{(3)}(M-1)^{-1}||\mathbf{h}_{n,k}||^2\sin^2\theta_{n,k}\right)\right]$$

$$\geq \mathbb{E}\left[\log_2\left(1 + \bar{S}_{n,k}^{(1)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2\right)\right] + \mathbb{E}\left[\log_2\left(1 + \frac{\bar{S}_{n,k}^{(3)}(M-1)^{-1}||\mathbf{h}_{n,k}||^2\sin^2\theta_{n,k}}{1 + \bar{S}_{n,k}^{(1)}M^{-1}||\mathbf{h}_{n,k}||^2}\right)\right]$$

$$-\mathbb{E}\left[\log_2\left(1 + \bar{S}_{n,k}^{(2)}M^{-1}||\mathbf{h}_{n,k}||^2 + \bar{S}_{n,k}^{(3)}(M-1)^{-1}||\mathbf{h}_{n,k}||^2\sin^2\theta_{n,k}\right)\right] \tag{54}$$

$$= \mathbb{E}\left[\log_2\left(1 + \bar{S}_{n,k}^{(1)}|\mathbf{h}_{n,k}\mathbf{w}_n|^2\right)\right] - \mathbb{E}\left[\log_2\left(1 + \bar{S}_{n,k}^{(1)}M^{-1}||\mathbf{h}_{n,k}||^2\right)\right]$$

$$+ \mathbb{E}\left[\log_2\left(1 + \frac{\frac{d_{n,k}^{-\alpha}P_n||\mathbf{h}_{n,k}||^2}{\sigma_{n,k}^2 \, KM}}{1 + \bar{S}_{n,k}^{(2)}M^{-1}||\mathbf{h}_{n,k}||^2 + \bar{S}_{n,k}^{(3)}(M-1)^{-1}||\mathbf{h}_{n,k}||^2\sin^2\theta_{n,k}}\right)\right] \tag{55}$$

$$\tilde{R}_{sum}^{LB}(1) = \sum_{n=1}^{N}\log_2\left(1 + \frac{\frac{d_{n,1}^{-\alpha}\phi_n P}{\sigma_{n,1}^2 K}}{\frac{M}{(M-1)} + \Gamma\left(\frac{2M-1}{M-1}\right)2^{-\frac{B}{NK(M-1)}}\left(\prod_{p=1}^{N}\prod_{q=1}^{K}\left(\frac{d_{p,q}^{-\alpha}}{\sigma_{p,q}^2}\right)^{\frac{1}{NK}}\right)C}\right) \tag{56}$$

$$\mathcal{L}_\phi(\phi_{n,k},\lambda_\phi) = -\prod_{n=1}^{N}\left(1 + \frac{\frac{d_{n,1}^{-\alpha}\phi_n P}{\sigma_{n,1}^2 K}}{\frac{M}{(M-1)} + \Gamma\left(\frac{2M-1}{M-1}\right)2^{-\frac{B}{NK(M-1)}}\left(\prod_{p=1}^{N}\prod_{q=1}^{K}\left(\frac{d_{p,q}^{-\alpha}}{\sigma_{p,q}^2}\right)^{\frac{1}{NK}}\right)C}\right) + \lambda_\phi\left(\sum_{n=1}^{N}\phi_n - 1\right) \tag{57}$$

$$\left.\frac{\partial\mathcal{L}_\phi(\phi_n,\lambda_\phi)}{\partial\phi_n}\right|_{\{\phi_n^\star\},\lambda_\phi^\star} = -\frac{\frac{d_{n,1}^{-\alpha}P}{\sigma_{n,1}^2 K}}{\frac{M}{(M-1)} + \Gamma\left(\frac{2M-1}{M-1}\right)\left(\prod_{p=1}^{N}\prod_{q=1}^{K}\left(\frac{d_{p,q}^{-\alpha}}{\sigma_{p,q}^2}\right)^{\frac{1}{NK}}\right)2^{-\frac{B}{NK(M-1)}}C}$$

$$\times \prod_{i=1,i\neq n}^{N}\left(1 + \frac{\frac{d_{i,1}^{-\alpha}P}{\sigma_{i,1}^2 K}\phi_i^\star}{\frac{M}{(M-1)} + \Gamma\left(\frac{2M-1}{M-1}\right)\left(\prod_{p=1}^{N}\prod_{q=1}^{K}\left(\frac{d_{p,q}^{-\alpha}}{\sigma_{p,q}^2}\right)^{\frac{1}{NK}}\right)2^{-\frac{B}{NK(M-1)}}C}\right) + \lambda_\phi^\star = 0 \tag{58}$$

$$\left.\frac{\partial\mathcal{L}_\phi(\phi_n,\lambda_\phi)}{\partial\lambda_\phi}\right|_{\{\phi_n^\star\}} = \sum_{n=1}^{N}\phi_n^\star - 1 = 0 \tag{59}$$

According to (58), we can obtain the following relationship for $\forall i \neq n$ that,

$$\phi_i^\star = \phi_n^\star + K\left(\frac{M}{(M-1)P} + \Gamma\left(\frac{2M-1}{M-1}\right)\right.$$

$$\left.\times \left(\prod_{p=1}^{N}\prod_{q=1}^{K}\left(\frac{d_{p,q}^{-\alpha}}{\sigma_{p,q}^2}\right)^{\frac{1}{NK}}\right)\frac{2^{-\frac{B}{NK(M-1)}}}{P}C\right)\left(\frac{\sigma_{n,1}^2}{d_{n,1}^{-\alpha}} - \frac{\sigma_{i,1}^2}{d_{i,1}^{-\alpha}}\right).$$

Substituting the above relationship into (59), we can obtain (21). Substituting $\phi_n^\star$ as a function of C given by (21) into the definition of C above, we can obtain the equation about the optimal $C^\star$ given by (22). Moreover, according to (2) and (3), it is easy to prove $\sum_{i=1,i\neq m}^{N}\left(\frac{\sigma_{m,1}^2}{d_{m,1}^{-\alpha}} - \frac{\sigma_{i,1}^2}{d_{i,1}^{-\alpha}}\right) \leq \sum_{i=1,i\neq n}^{N}\left(\frac{\sigma_{n,1}^2}{d_{n,1}^{-\alpha}} - \frac{\sigma_{i,1}^2}{d_{i,1}^{-\alpha}}\right)$, $\forall 1 < m < n < N$. Then, the order of the coefficients ϕ_is follows.

REFERENCES

[1] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, Dresden, Germany, Jun. 2013, pp. 1–5.

[2] B. Kim *et al.*, "Non-orthogonal multiple access in a downlink multiuser beamforming system," in *Proc. IEEE Mil. Commun. Conf.*, San Diego, CA, USA, Nov. 2013, pp. 1278–1283.

[3] 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.

[4] 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.

[5] J. Choi, "On the power allocation for a practical multiuser superposition scheme in NOMA systems," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 438–441, Mar. 2016.

[6] C.-L. Wang, J.-Y. Chen, and Y.-J. Chen, "Power allocation for a downlink non-orthogonal multiple access system," *IEEE Wireless Commun. Lett.*, vol. 5, no. 5, pp. 532–535, Oct. 2016.

[7] Z. Ding *et al.*, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[8] Y. Zhang, H.-M. Wang, T.-X. Zheng, and Q. Yang, "Energy-efficient transmission design in non-orthogonal multiple access," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2852–2857, Mar. 2017.

[9] J. Cui, Z. Ding, and P. Fan, "Outage probability constrained MIMO-NOMA designs under imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8239–8255, Dec. 2018.

[10] J. Choi, "On the power allocation for MIMO-NOMA systems with layered transmissions," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3226–3237, May 2016.

[11] 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.

[12] Y. I. Choi, J. W. Lee, M. Rim, and C. G. Kang, "On the performance of beam division nonorthogonal multiple access for FDD-based large-scale multi-user MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5077–5089, Aug. 2017.

[13] Z. Chen, Z. Ding, P. Xu, and X. Dai, "Optimal precoding for a QoS optimization problem in two-user MISO-NOMA downlink," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1263–1266, Jun. 2016.

[14] Z. Chen, Z. Ding, X. Dai, and G. K. Karagiannidis, "On the application of quasi-degradation to MISO-NOMA downlink," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6174–6189, Dec. 2016.

[15] V. D. Nguyen, H. D. Tuan, T. Q. Duong, H. V. Poor, and O.-S. Shin, "Precoder design for signal superposition in MIMO-NOMA multicell networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2681–2695, Dec. 2017.

[16] C. Chen, W. Cai, X. Cheng, L. Yang, and Y. Jin, "Low complexity beamforming and user selection schemes for 5G MIMO-NOMA systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2708–2722, Dec. 2017.

[17] Q. Zhang, Q. Li, and J. Qin, "Robust beamforming for nonorthogonal multiple-access systems in MISO channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10231–10236, Dec. 2016.

[18] Z. Ding and H. V. Poor, "Design of massive-MIMO-NOMA with limited feedback," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 629–633, May 2016.

[19] S. Liu and C. Zhang, "Non-orthogonal multiple access in a downlink multiuser beamforming system with limited CSI feedback," *EURASIP J. Wireless Commun. Netw.*, vol. 2016, no. 1, p. 239, Oct. 2016.

[20] Q. Yang, H.-M. Wang, D. W. K. Ng, and M. H. Lee, "NOMA in downlink SDMA with limited feedback: Performance analysis and optimization," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2281–2294, Oct. 2017.

[21] X. Chen, Z. Zhang, C. Zhong, and D. W. K. Ng, "Exploiting multiple-antenna techniques for non-orthogonal multiple access," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2207–2220, Oct. 2017.

[22] L. Zhang, J. Liu, M. Xiao, G. Wu, Y.-C. Liang, and S. Li, "Performance analysis and optimization in downlink NOMA systems with cooperative full-duplex relaying," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2398–2412, Oct. 2017.

[23] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[24] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1478–1491, Sep. 2007.

[25] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation*, document TS 36.211 Version. 15.5.0, 3GPP, 2018.

[26] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York, NY, USA: Academic, 2007.

[27] A. Lew and H. Mauch, *Dynamic Programming: A Computational Tool*. Berlin, Germany: Springer, 2006.

[28] S. V. Amari and R. B. Misra, "Closed-from expressions for distribution of sum of exponential random variables," *IEEE Trans. Rel.*, vol. 46, no. 4, pp. 519–522, Dec. 1997.

[29] M.-S. Alouini and A. J. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1165–1181, Jul. 1999.

[30] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, vol. 55. Chelmsford, MA, USA: Courier Corporation, 1965.
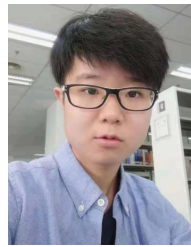
Zhiyao Tang received the bachelor's degree from Xidian University, Xi'an, China, in 2017. She is currently pursuing the master's degree with the School of Electronic and Information Engineering, Beihang University, Beijing, China. Her research interests focus on physical-layer security, as well as non-orthogonal multiple-access and MIMO systems.
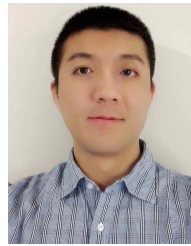
Liang Sun (S'08–M'12) received the Ph.D. degree in electrical engineering from The Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2010.

From 2010 to 2013, he was a Researcher with the Alcatel-Lucent Shanghai Bell and the NEC Corporation. From September 2014 to December 2015, he was a Post-Doctoral Research Fellow at the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada. Since 2016, he has been with Beihang University as an Associate Professor. He also held visiting positions at the Waseda University, Japan, and the South University of Science and Technology of China, Shenzhen, China. His research interests include information theory and signal processing; in particular design and analysis of wireless communication systems, MIMO, physical-layer security, ultra-dense heterogenous networks, the Internet of Things, and limited feedback techniques. He received the 2010 Young Author Best Paper Award by the IEEE Signal Processing Society.

Lu Cao received the bachelor's degree from the Dalian University of Technology, Dalian, China, in 2017. She is currently pursuing the master's degree with the School of Electronic and Information Engineering, Beihang University, Beijing, China. Her research interests focus on physical-layer security, non-orthogonal multiple-access, MIMO systems, and ultra-dense heterogenous wireless networks.
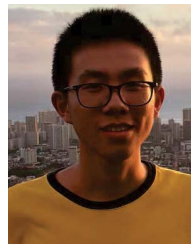
Shutong Qi is currently a junior undergraduate student with the School of Electronic and Information Engineering, Beihang University, Beijing, China. In Summer 2019, he was a Research Assistant with the Computer Science Department, Dartmouth College, Hanover, NH, USA. His research interests include electromagnetic, MIMO systems, and human–computer interaction.

Yong Feng (M'15) received the Ph.D. degree in information and communication engineering from the University of Electronic Science and Technology of China in 2011. He is currently an Associate Professor with the Yunnan Key Laboratory of Computer Technology Applications, Kunming University of Science and Technology, China. His current research interests include wireless sensor networks, vehicular ad hoc networks, and the Internet of Things. He has authored or coauthored three books and over 40 papers in peer-reviewed journals, and conferences.