Optimal Power Allocation for the Downlink of Cache-aided NOMA Systems

Yaru Fu*, Hong Wang†, Chi Wan Sung*
*Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR
†Department of Communication Engineering, Nanjing University of Posts & Telecommunications, China
Email: yarufu2-c@my.cityu.edu.hk, wanghong@njupt.edu.cn, albert.sung@cityu.edu.hk

***Abstract*—In this work, we investigate the optimal power allocation for the downlink of cache-aided NOMA systems, in which the user with weaker link condition caches the intended message of user with stronger channel condition within each cell, resulting in different signal decoding at receiver side from that of conventional NOMA. The goal is to minimize the total needed power of all users subject to the least data rate constraint of each user. To solve this non-convex optimization problem, we first transform the required power of users in each cell to that of their serving BS. With the power constraint transformation, the original power allocation problem can be transformed into a power optimization problem with reduced dimensions. The optimality of the newly formulated problem is analyzed. In addition, we show that the optimal solutions to the two optimization problems are unique and one-to-one mapping, which motivates us to get the optimal power control for the original problem by solving the newly formulated problem. An iterative power control algorithm is then designed, which is guarantee to converge to the unique optimal point given that the feasible power region is non-empty. Simulation results demonstrate the power-saving capability of our proposed power allocation scheme for cache-aided NOMA.**

***Index Terms*—NOMA, caching, SIC, optimal power allocation.**

I. Introduction

Non-orthogonal multiple access (NOMA) has been treated as a promising candidate for the fifth generation of wireless communication systems (5G) owning to its capability to meet the dramatically increased data rate requirements [1]. The main existing NOMA schemes can be classified as two categories: code-domain NOMA and power-domain NOMA. In code-domain NOMA, multiple users are identified by non-orthogonal chip codes. Meanwhile, in power-domain NOMA, different users with diverse channel conditions are allowed to access the same time-frequency resource simultaneously. In this paper, the power control for power-domain NOMA is considered. Without otherwise stated, we adopt NOMA to indicate power-domain NOMA.

Since successive interference cancellation (SIC) should be implemented at the multiplexed users, power control for the users becomes indispensable as it is the key to do interference elimination and signal decoding [2]. In [3], [4], fractional transmit power control (FTPC) was adopted, which is a sub-optimal power allocation strategy and assigns power in accordance with users' channel conditions. The optimal power control for NOMA networks was investigated by [5], [6], and it was demonstrated that the optimal scheme requires less power than FTPC and outperforms its orthogonal multiple access (OMA) counterparts.

From a technical perspective, caching is another technique to enhance the spectral efficiency of 5G wireless networks [7], [8]. In a physical layer caching based NOMA system, the user with weaker channel condition caches the prior information of stronger user such that the interference elimination can be adopted at both the strong and weak users. In [9], the performance enhancement of caching based NOMA when compared to that of conventional NOMA in terms of outage probability is investigated. Besides, it is shown in [10] that caching can help to improve the converge probability and the spectral efficiency of 5G NOMA networks. However, the aforementioned caching aided NOMA works focus on single-cell scenario. Optimal power allocation for caching based NOMA system with multiple cells has not been discussed to the best of our knowledge.

In this work, the optimal power allocation for a caching assisted NOMA system with two interactional cells is studied. We target at minimizing the total transmit power under the data rate constraint of each user. A distributed scheme is proposed to obtain the optimal power control. First, we derive the needed transmission power of each BS to meet its associated users' data rate requirements. With the transformation of power constraints, the original power control problem can be transformed into a new power minimization problem for BSs. Then, we analyze the optimality of the newly formulated problem and its equivalence to the original optimization problem. Subsequently, an iterative algorithm is designed to obtain the optimal power allocation. Simulation results show the validity of our proposed method in terms of convergence and power consumption.

The organization of the rest parts of this paper is given as follows. In Section II, the system model and problem formulation is stated. In Section III, we transform the power

This paper was supported by a grant from the Research Grants Council of the Hong Kong SAR, China (Project No. CityU 11216416), a grant of the Natural Science Foundation of China (Project No. 61801246), and a grant of the Natural Science Foundation of Jiangsu Province (Project No. BK20170910).

978-1-5386-6119-2/18/$31.00 ©2018 IEEE

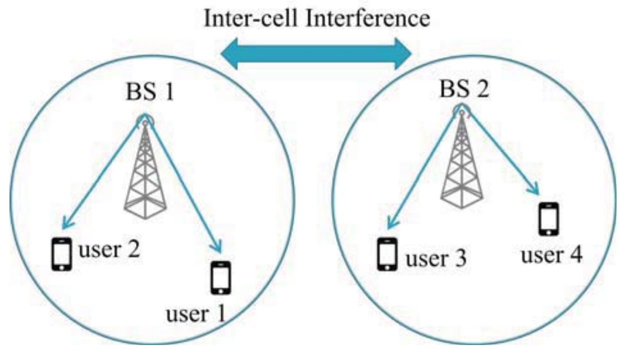

Fig. 1. Caching based downlink NOMA system model

constraints and reformulate the original problem as a new power minimization problem for BSs. Additionally, the optimality of the newly formulated problem and its relationship to the original problem are analyzed. Section IV describes the proposed iterative power allocation algorithm and its provable convergence guarantee. Simulation results are presented in Section V. Finally, we summarize the main work of this paper in Section VI.

II. System Model and Problem Formulation

In this section, we first introduce the system model to be studied in this work and then formulate the power minimization problem mathematically.

A. System Model

The downlink scenario of two-cell caching based NOMA systems is considered. Two cells are labeled as cell 1 and cell 2. In each cell, there is one BS located at the center of the cell and serving two users[1]. Let user 1 and user 2 be served by BS 1, while assume users 3 and 4 are associated with cell 2. The system model of two-cell cache-assisted NOMA is shown in Fig. 1. Denote by $\mathcal{I}=\{1,2,3,4\}$ and $\mathcal{J}=\{1,2\}$ the sets of indexes of users and cells, respectively. For $i\in\mathcal{I}$, define g_i as the channel gain between user i and its associated BS. In addition, let $g_{i,j}$ be the interfering channel gain between BS j and user i, where $j\in\mathcal{J}$, $i\in\mathcal{I}$. Furthermore, denote by η the noise power at each user.

The cache aided system is investigated in this work. Specifically, within each cell, the user with weaker channel condition caches the information of the other user [9]. Without loss of generality, we assume $g_1<g_2$ and $g_3<g_4$. Therefore, user 1 caches the information of user 2, meanwhile, user 3 has the prior information of user 4. In downlink NOMA system, each BS transmits the superposition code of its attached users' messages. For $i\in\mathcal{I}$, define p_i as the transmission power of user i. Denote by $q_1=p_1+p_2$ and $q_2=p_3+p_4$ the total transmit power of BSs 1 and 2, respectively. Let $\boldsymbol{p}=(p_1,p_2,p_3,p_4)$. Besides, denote by r_i the required data rate of user i. Based on the central limit theorem [11], the summation of a large number of random variables tends to follow a distribution named Gaussian. Therefore, the inter-cell interference is assumed to be additive white Gaussian noise in this work. Define I_i as the normalized inter-cell interference plus noise power of user i, and they are expressed as

$$\begin{aligned} I_i &= q_2\tilde{g}_{i,2}+\tilde{\eta}_i, i=1,2,\\ I_u &= q_1\tilde{g}_{u,1}+\tilde{\eta}_u, u=3,4, \end{aligned} \tag{1}$$

in which $\tilde{g}_{i,j}=g_{i,j}/g_i$ and $\tilde{\eta}_i=\eta/g_i$ for $i\in\mathcal{I}$ and $j\in\mathcal{J}$.

We apply a binary variable, $a_1(q_2)$, to indicate the relationship between I_1 and I_2 in cell 1. In detail, $a_1(q_2)=1$ if $I_1\geq I_2$, and 0 otherwise. Similarly, we define $a_2(q_1)$ for cell 2, i.e., $a_2(q_1)=1$ if $I_3\geq I_4$, and $a_2(q_1)=0$ otherwise.

With aforementioned definitions, the signal decoding in the caching aided NOMA system is then considered. We take cell 1 as an example. When $a_1(q_2)=1$, the capacity of user 2 can be obtained by performing successive interference cancellation (SIC), which removes the interference from user 1 before decoding its intended message. In contrast to conventional NOMA system, user 1 can also subtract the signal of user 2 first and then decode its desired message as it caches the prior information of user 2. Mathematically, when $a_1(q_2)=1$, the data rates of user 1 and 2 are given by

$$\begin{aligned} \hat{R}_1(\tilde{a}_1,\boldsymbol{p}) &= W\log_2(1+\frac{p_1}{I_1}),\ \tilde{a}_1=1,\\ \tilde{R}_2(\tilde{a}_1,\boldsymbol{p}) &= W\log_2(1+\frac{p_2}{I_2}),\ \tilde{a}_1=1, \end{aligned} \tag{2}$$

where W represents the system bandwidth.

In addition, the capacity of user 2 to decode the signal of user 1 when performing SIC is expressed as

$$\hat{R}_{1,2}(\tilde{a}_1,\boldsymbol{p}) = W\log_2(1+\frac{p_1}{p_2+I_2}),\ \tilde{a}_1=1. \tag{3}$$

Note that in conventional NOMA networks, $\hat{R}_{1,2}(1,\boldsymbol{p}) > \hat{R}_1(1,\boldsymbol{p})$ is always true when $I_1>I_2$. However, this relationship may not hold for cache assisted NOMA system since user 1 has the prior information of user 2.

For the scenario where $a_1(q_2)=0$, user 2 decodes its signal by treating the signal of user 1 as noise. Meanwhile, user 1 can remove the interference from user 2 by using its cached prior information or via performing SIC. The obtained data rates are given as follows:

$$\begin{aligned} \tilde{R}_1(\tilde{a}_1,\boldsymbol{p}) &= W\log_2(1+\frac{p_1}{I_1}),\ \tilde{a}_1=0,\\ \tilde{R}_2(\tilde{a}_1,\boldsymbol{p}) &= W\log_2(1+\frac{p_2}{p_1+I_2}),\ \tilde{a}_1=0. \end{aligned} \tag{4}$$

For notation clearly, we write two auxiliary function $\tilde{R}_i(\tilde{a}_1,\boldsymbol{p})$ for $i\in\{1,2\}$ as follows

$$\tilde{R}_1(\tilde{a}_1,\boldsymbol{p}) = \begin{cases} \min\{W\log_2(1+\frac{p_1}{I_1}), W\log_2(1+\frac{p_1}{p_2+I_2})\} & \text{if } \tilde{a}_1=1,\\ W\log_2(1+\frac{p_1}{I_1}) & \text{if } \tilde{a}_1=0, \end{cases} \tag{5}$$

[1]In practical, there will be an arbitrary number of users in each cell and user clustering can be adopted to separate users into different groups where each group is scheduled on the orthogonal resource. In this work, we just assume each cell has one such cluster. The joint user grouping and power control for the general scenario will be considered in the future work.

and

$$\tilde{R}_2(\tilde{a}_1, \boldsymbol{p}) = \begin{cases} W\log_2(1+\frac{p_2}{I_2}) & \text{if } \tilde{a}_1 = 1, \\ W\log_2(1+\frac{p_2}{p_1+I_2}) & \text{if } \tilde{a}_1 = 0. \end{cases} \tag{6}$$

Define $R_i(\boldsymbol{p})$ as the obtainable transmit data rate of user i. Therefore, we have

$$R_i(\boldsymbol{p}) = \tilde{R}_i(a_1(q_2), \boldsymbol{p}), \text{ for } i \in \{1, 2\}. \tag{7}$$

Similarly, the data rates of user 3 and user 4 can be given as follows:

$$R_u(\boldsymbol{p}) = \tilde{R}_u(a_2(q_1), \boldsymbol{p}), \text{ for } u \in \{3, 4\}, \tag{8}$$

in which

$$\tilde{R}_3(\tilde{a}_2, \boldsymbol{p}) = \begin{cases} \min\{W\log_2(1+\frac{p_3}{I_3}), W\log_2(1+\frac{p_3}{p_4+I_4})\} & \text{if } \tilde{a}_2 = 1, \\ W\log_2(1+\frac{p_3}{I_3}) & \text{if } \tilde{a}_2 = 0, \end{cases} \tag{9}$$

and

$$\tilde{R}_4(\tilde{a}_2, \boldsymbol{p}) = \begin{cases} W\log_2(1+\frac{p_4}{I_4}) & \text{if } \tilde{a}_2 = 1, \\ W\log_2(1+\frac{p_4}{p_3+I_4}) & \text{if } \tilde{a}_2 = 0. \end{cases} \tag{10}$$

B. Problem Formulation

In this work, we target at minimizing the total transmit power of all users under each user's data rate constraint. Mathematically, the problem formulation is given as follows:

$$\min \sum_{i=1}^{4} p_i \tag{11}$$

subject to

$$\begin{aligned} &C1: \ R_i(\boldsymbol{p}) \geq r_i, \ \ i \in \mathcal{I}, \\ &C2: \ p_i \geq 0, \ i \in \mathcal{I}, \end{aligned} \tag{12}$$

where $C1$ represents the minimum data rate requirement of each user. In addition, $C2$ indicates the nonnegativity of the allocated power to each user. We denote by $\mathcal{P} \subseteq \mathbb{R}_+^4$ the feasible power region of problem (11), where $\mathbb{R}_+^4$ is the set of four-dimensional non-negative real vectors. Note that $\boldsymbol{p}$ is feasible if and only if $C1$ and $C2$ are met at $\boldsymbol{p}$.

III. Required Power at a BS and a Related Power Control Problem

In this section, we first investigate the needed transmission power of a BS to satisfy the data rate requirements of its attached users. Based on the power requirement at BSs, we formulate an equivalent optimization problem to problem (11)-(12) with reduced dimensions. Subsequently, the optimality of the new-formulated problem and its relationship to that of the original problem are analyzed.

A. Required Power at a BS

For $j \in \mathcal{J}$, we define $f_j(q_{-j})$ as the least needed power of BS j to meet its attached users' data rate constraints. In addition, let $\gamma_i = 2^{r_i/W} - 1$ for $i \in \mathcal{I}$. We consider BS 1 firstly. Suppose the transmit power of users 3 and 4 are fixed, i.e., q_2 is given and we want to meet the rate constraints of users in cell 1 given as

$$\begin{aligned} &\tilde{R}_1(\tilde{a}_1, \boldsymbol{p}) \geq r_1, \\ &\tilde{R}_2(\tilde{a}_2, \boldsymbol{p}) \geq r_2. \end{aligned} \tag{13}$$

It's obvious that the minimum required power of user 1 and user 2 is obtained by letting the inequalities in (13) hold with equalities. Therefore, we have the following definition:

Definition 1. *The needed transmit power of BS 1 equals to the sum of user 1 and user 2's transmission power, which is given by*

$$\tilde{f}_1(\tilde{a}_1, q_2) = \begin{cases} \max\{\gamma_1 I_1 + \gamma_2 I_2, \gamma_1\gamma_2 I_2 + \gamma_1 I_2 + \gamma_2 I_2\} & \textit{if } \tilde{a}_1 = 1, \\ \gamma_1 I_1 + \gamma_2 I_2 + \gamma_1\gamma_2 I_1 & \textit{if } \tilde{a}_1 = 0. \end{cases} \tag{14}$$

Similarly, once $q_1 \geq 0$ and $\tilde{a}_2$ are given, the required power of BS 2 is obtained by Definition 2:

Definition 2. *The required power of BS 2 to satisfy its associated users' data rate requirements is given by*

$$\tilde{f}_2(\tilde{a}_2, q_1) = \begin{cases} \max\{\gamma_3 I_3 + \gamma_4 I_4, \gamma_3\gamma_4 I_4 + \gamma_3 I_4 + \gamma_4 I_4\} & \textit{if } \tilde{a}_2 = 1, \\ \gamma_3 I_3 + \gamma_4 I_4 + \gamma_3\gamma_4 I_3 & \textit{if } \tilde{a}_2 = 0. \end{cases} \tag{15}$$

With aforementioned definitions, the following Lemma can be obtained:

Lemma 3. *For $j \in \mathcal{J}$, the least required transmit power at BS j should satisfy*

$$f_j(q_{-j}) = \min\{\tilde{f}_j(0, q_{-j}), \tilde{f}_j(1, q_{-j})\}. \tag{16}$$

Proof: With aforementioned analysis, when $j = 1$ and q_2 is fixed, the value of $f_1(q_2)$ is either $\tilde{f}_1(1, q_2)$ or $\tilde{f}_1(0, q_2)$. When $a_1(q_2) = 1$, it is easy to obtain that

$$\begin{aligned} &\tilde{f}_1(1, q_2) - \tilde{f}_1(0, q_2) \\ &= \max\{-\gamma_1\gamma_2 I_1, -\gamma_1(1+\gamma_2)(I_1 - I_2)\} < 0. \end{aligned} \tag{17}$$

In addition, when $a_1(q_2) = 0$, we have $\tilde{f}_1(1, q_2) > \tilde{f}_1(0, q_2)$. Therefore, (16) holds when $j = 1$. Through a similar proof technique, we can get that (16) is true when $j = 2$. □

B. A Related Power Optimization Problem

In this subsection, we formulate a power optimization problem for BSs with the aforementioned power constraints transformation.

The optimization problem aims at minimizing the total power of all BSs with the consideration of each BS's power constraint. Mathematically, it is presented as

$$\min \sum_{j \in \mathcal{J}} q_j \tag{18}$$

subject to

$$q_j \geq f_j(q_{-j}),\ j \in \mathcal{J}, \tag{19}$$

$$q_j \geq 0,\ j \in \mathcal{J}, \tag{20}$$

where f_j in (19) is well defined by Lemma 3. In addition, (20) indicates the nonnegativity of BS's transmission power. We define its feasible power region as $\mathcal{Q} \subseteq \mathbb{R}_+^2$, in which $\mathbb{R}_+^2$ is the set of two-dimensional non-negative real vectors.

Lemma 4. *If the optimal solution to problem* (18) *exists, it must be achieved when the inequality in* (19) *holds with equality.*

Proof: Define (q_1^*, q_2^*) as the optimal solution, and assume there exists one BS meeting the power constraint with strict inequality. Let us suppose BS 1 is such a BS without loss of generality, i.e., $q_1^* > f_1(q_2^*)$. We reduce q_1^* to $\hat{q}_1$ such that $\hat{q}_1 = f_1(q_2^*)$ with $\hat{q}_1 < q_1^*$. It is obvious that $(\hat{q}_1, q_2^*)$ meets the requirements in (19), which contradicts that (q_1^*, q_2^*) is the optimal solution. □

We consider the following power control subproblem with parameters $\tilde{a}_1$ and $\tilde{a}_2$:

$$\min \sum_{j \in \mathcal{J}} q_j \tag{21}$$

subject to

$$q_j \geq \tilde{f}_j(\tilde{a}_j, q_{-j}),\ j \in \mathcal{J}, \tag{22}$$

$$q_j \geq 0,\ j \in \mathcal{J}. \tag{23}$$

Since both $\tilde{a}_1$ and $\tilde{a}_2$ can be 1 or 0, there are four such power control subproblems in total.

Theorem 5. *If problem* (18) *is feasible, its optimal solution is unique.*

Proof: If problem (18) is feasible, some of the power control subproblems in (21) must be feasible, and each power control subproblem has a unique optimal solution [12]. Among them, one needs the minimum transmission power is the optimal value to problem (18).

For the uniqueness of optimal solution to problem (18), let $\boldsymbol{q}^* = (q_1^*, q_2^*)$ and $\hat{\boldsymbol{q}} = (\hat{q}_1, \hat{q}_2)$ be two different optimal solutions. Let $q_1^* > \hat{q}_1$ without loss of generality. Based on the definition of $f_2(q_1)$, we have $f_2(q_1^*) > f_2(\hat{q}_1)$. According to Lemma 4, we have $q_2^* = f_2(q_1^*) > f_2(\hat{q}_1) = \hat{q}_2$. As a result, $q_1^* + q_2^* > \hat{q}_1 + \hat{q}_2$, making a contradiction. □

Before analyzing the relationship between the optimal solution to problem (11) and problem (18), a mapping Φ from domain $\mathcal{Q}$ to co-domain $\mathcal{P}$ is defined. For any $(q_1, q_2) \in \mathcal{Q}$, we must have $q_j \geq f_j(q_{-j})$ for $j \in \mathcal{J}$. With fixed q_2, via letting the data rate requirements of user 1 and user 2 hold with equality, we can find $\hat{p}_1$ and $\hat{p}_2$ satisfying $\hat{p}_1 + \hat{p}_2 = f_1(q_2) \leq q_1$ as well as the least data rate constraints of users 1 and 2. Let $(p_1, p_2) = \beta(\hat{p}_1, \hat{p}_2)$ with $\beta = \frac{q_1}{\hat{p}_1 + \hat{p}_2} \geq 1$. Obviously, (p_1, p_2) also satisfies the rate requirements of users 1 and 2. Similarly, we can define (p_3, p_4). Let $\boldsymbol{q} = (q_1, q_2)$. The mapping Φ is defined as: $\Phi(\boldsymbol{q}) = (p_1, p_2, p_3, p_4) \in \mathcal{P}$. Denote by $\hat{\mathcal{P}} \subseteq \mathcal{P}$ the region of Φ. In addition, another mapping $\Psi : \mathcal{P} \rightarrow \mathcal{Q}$ is defined as follows: $\Psi(p_1, p_2, p_3, p_4) = (p_1 + p_2, p_3 + p_4)$. It is obvious that $(p_1 + p_2, p_3 + p_4) \in \mathcal{Q}$ if $(p_1, p_2, p_3, p_4) \in \mathcal{P}$. Note that we restrict the domain of Ψ to be $\hat{\mathcal{P}}$. Therefore, Ψ can be regarded as the inverse function of Φ.

Lemma 6. *If problem* (11) *gets the optimal solution, it satisfies the inequalities in $C1$ with equalities.*

The proof of this Lemma is similar to that of Lemma 4. Details are omitted here.

Lemma 7. *Problems* (11) *and* (18) *have the same optimal value.*

Proof: Denote by $\boldsymbol{p}^*$ and $\boldsymbol{q}^*$ the optimal power allocation to problems (11) and (18), respectively. It is easy to check that

$$||\boldsymbol{q}^*||_1 = ||\Phi(\boldsymbol{q}^*)||_1 \geq ||\boldsymbol{p}^*||_1 = ||\Psi(\boldsymbol{p}^*)||_1 \geq ||\boldsymbol{q}^*||_1$$

in which $||\boldsymbol{x}||_1$ represents the l_1-norm of vector $\boldsymbol{x}$, which equals to the sum of each component in $\boldsymbol{x}$. □

Theorem 8. *If the feasible power region of problem* (11) *is non-empty, its optimal solution $\boldsymbol{p}^*$ is unique, and $\boldsymbol{p}^* = \Phi(\boldsymbol{q}^*)$, in which $\boldsymbol{q}^*$ represents the optimal solution to problem* (18).

Proof: According to the definitions of Φ and Ψ, if problem (11) is feasible, problem (18) is also feasible. Based on Theorem 5, the optimal solution to problem (18) is unique and we denote it by $\boldsymbol{q}^*$. In accordance with Lemma 7, the optimal values of problems (11) and (18) are the same, which means $\boldsymbol{p}^* = (p_1^*, p_2^*, p_3^*, p_4^*) = \Phi(\boldsymbol{q}^*)$ should be an optimal power vector.

Let $\hat{\boldsymbol{p}} = (\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4)$ be a different optimal solution to problem (11). According to Lemma 7, $\Psi(\hat{\boldsymbol{p}})$ should be the optimal solution to (18) and equals to $\boldsymbol{q}^*$, which means $q_1^* = \hat{p}_1 + \hat{p}_2 = p_1^* + p_2^*$ and $q_2^* = \hat{p}_3 + \hat{p}_4 = p_3^* + p_4^*$. Based on Lemma 6, the optimal solution is achieved when the inequality rate constraints are satisfied with equality. With fixed q_2^*, there must exist $\hat{p}_1 = p_1^*$ and $\hat{p}_2 = p_2^*$. Similarly, we have $\hat{p}_3 = p_3^*$ and $\hat{p}_4 = p_4^*$. □

Theorem 8 inspires us to get the optimal solution of problem (11) via solving the corresponding problem (18).

IV. Iterative Power Allocation Algorithm

In this section, we design an iterative power allocation strategy for problem (18). Given that the feasible power region is non-empty, the algorithm is ensured to converge since our algorithm falls into Yate's power control framework [13].

For $j \in \mathcal{J}$, denote by $q_j^{(t)}$ the total transmission power of BS j at t-th iteration. At iteration t, two BSs update their required power to satisfy the rate requirements of its associated users, which is given by

$$q_j^{(t)} = f_j(q_{-j}^{(t-1)}),\ j \in \mathcal{J}. \tag{24}$$

According to [13], this iterative power allocation method can be written into a vector form as follows:

$$\boldsymbol{q}^{(t)} = \mathbb{F}(\boldsymbol{q}^{(t-1)}), \tag{25}$$

in which $\mathbb{F}$ is the interference function, whose j-th component can be regarded as the effective interference that BS j suffers from. For our system,

$$\mathbb{F}_1(\boldsymbol{q}) = f_1(q_2) = \min\{\tilde{f}_1(1, q_2), \tilde{f}_1(0, q_2)\}, \tag{26}$$

$$\mathbb{F}_2(\boldsymbol{q}) = f_2(q_1) = \min\{\tilde{f}_2(1, q_1), \tilde{f}_2(0, q_1)\}. \tag{27}$$

In addition, we say $\mathbb{F}$ is standard once it meets the following three criteria for any $\boldsymbol{q} \in \mathbb{R}_+^2$:

1) $\mathbb{F}(\boldsymbol{q}) \geq 0$.
2) If $\boldsymbol{q} \geq \hat{\boldsymbol{q}}$, then $\mathbb{F}(\boldsymbol{q}) \geq \mathbb{F}(\hat{\boldsymbol{q}})$.
3) For all $\alpha > 1$, then $\alpha\mathbb{F}(\boldsymbol{q}) \geq \mathbb{F}(\alpha\boldsymbol{q})$.

The following theorem shows the convergence of the proposed iterative power control algorithm.

Theorem 9. *Assume that the optimization problem* (11) *is feasible, the proposed iterative power allocation method in* (25) *is guarantee to converge to the unique optimal solution with an arbitrary initial power vector* $\boldsymbol{q}^{(0)}$.

Proof: If the feasible region of problem (11) is non-empty, by Theorem 5 and Theorem 8, (11) has a unique optimal solution. Based on Lemma 4, the optimal solution is obtained when the inequality constraints in (19) hold with equalities, which means the optimal solution is a fixed point.

For $a_1 = 1$, it can be seen from (14) that the two items in $\tilde{f}_1(1, q_2^{(t-1)})$ are affine functions to q_2. Each of the items meets three criteria of standard. By [13], the maximization operation preserves the property of standard. Therefore, $\tilde{f}_1(1, q_2^{(t-1)})$ is standard. When $a_1 = 0$, $\tilde{f}_1(0, q_2^{(t-1)})$ is also an affine function, and thus standard. The minimum operation also keeps the property of standard [13], thus, $\mathbb{F}_1(\boldsymbol{q})$ is standard. Through applying a similar proof technique, we show that $\mathbb{F}_2(\boldsymbol{q})$ is standard. This completes the proof that $\mathbb{F}(\boldsymbol{q})$ is standard.

Since the interference function of our power allocation strategy is standard, by [13], the algorithm can be guaranteed to converge to the unique fixed point. As aforementioned, the optimal solution is a fixed point, thus, the algorithm converges to the unique optimal solution. □

Once the optimal solution to problem (18) is obtained by the iterative power allocation method, the optimal solution to problem (11) can be obtained by applying Φ to $\boldsymbol{q}^*$ according to Theorem 8.

Remark 10. *The proposed power allocation method can be implemented in a distributed manner, which means each BS only requires some local information, i.e., the channel gains and the interference plus noise power of its associated users. Information interaction between BSs is not needed, making the algorithm attractive for practical application. That is why we do the problem transformation and propose the iterative power control strategy.*

V. Simulation Results

In this section, we use Monte-carlo simulation to demonstrate the performance of the designed power allocation algorithm for caching based NOMA networks. The radius of each cell is assumed to be $R = 500$ m. For each cell, the BS is located at the cell center and two users are uniformly distributed within it. The distance between two BSs, R_d, is set to 1000 m. The system bandwidth W is assumed to be 5 MHz and the noise power spectral density N_0 is set to -174 dBm/Hz. In addition, we consider the propagation model with both large-scale fading and small-scale fading [14]. Specifically, the large-scale path loss is distance-dependent, and it is given by $128.1 + 37.6\log_{10} d$, where d represents the distance between BS and user in km. For small-scale fading, each user is assumed to experience an independent Rayleigh fading. Without loss of generality, let the minimum rate requirement of user i be $r_i = r$ for $i \in \mathcal{I}$. The details of simulation parameters are summarized in Table I.

TABLE I
Computer simulation parameters

Parameters	Value
Cell radius, R	500 m
Distance between two BSs, R_d	1000 m
Large-scale fading	$128.1 + 37.6\log_{10} d$ dB, d is in km
Small-scale fading	Rayleigh fading
Users distribution scheme	Randomly uniform distribution
Noise power spectral density	-174 dBm/Hz
Overall system bandwidth, W	5 MHz
Throughput calculation	Shannon's capacity formula
Iteration terminal parameter	10^{-5}
Data rate requirement r	[5, 10, 15, 20, 25] Mbits/s

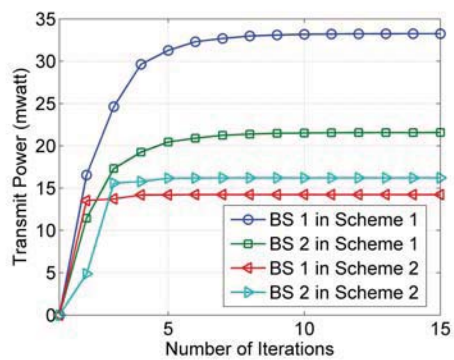

Fig. 2. Convergence performance

We adopt two information caching schemes in the simulation:

- Scheme 1: Within each cell, we randomly choose a user and assume it caches the message of the other user;
- Scheme 2: In each cell, the weak user caches the prior information of the strong user.

Two indicators are considered when evaluating the performance of our proposed power control algorithm for caching aided NOMA, i.e., 1) convergence performance, and 2) energy-saving ratio.

A. Convergence Analysis

Fig. 2 shows the convergence of the designed iterative power allocation method, in which the data rate requirement r is set to be 10 Mbits/s and the initial transmission power of each BS is assumed to be 0. We randomly generate a feasible instance for

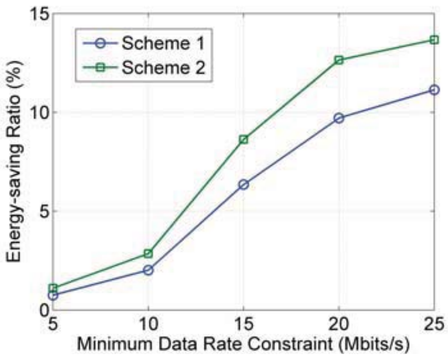

Fig. 3. Energy saving versus data rate requirements, r

each caching scheme and plot the transmit power of the BSs during iterations to demonstrate the convergence performance. The x-axis indicates the number of iterations, meanwhile the y-axis represents the transmit power. It is easy to see that the iterative power control algorithm takes only several steps to converge.

B. Energy-saving Ratio

Fig. 3 indicates the power-saving capability of the proposed power allocation for caching based NOMA when compared to that of the conventional power-controlled NOMA [5]. Since it was demonstrated by [5] and [15] that power-controlled NOMA requires less energy than OMA counterparts, we only compare the performance with NOMA in this subsection. The x-axis and y-axis are the minimum data rate requirement of each user and the energy-saving ratio, respectively. Specifically, the computational method of energy-saving ratio is given as follows:

$$\frac{|P_N - P_{CN}|}{P_{CN}} \times 100\%, \tag{28}$$

in which P_N and P_{CN} respectively refer to the needed transmit power of power-controlled NOMA [5] and the caching aided NOMA that applies the proposed optimal power allocation strategy. We consider two aforementioned caching schemes. It can be seen that for each caching scheme, the cache-aided NOMA is more energy efficient than conventional NOMA, particularly when each user has a high data rate requirement. For example, when r is set to be 25 Mbits/s, caching based NOMA can respectively reduce power by 11.13% and 13.67% under Scheme 1 and Scheme 2 when compared to that of power-controlled NOMA. Besides, with the increasing of data rate requirement, the energy-saving ratios of both schemes increase. In addition, for any given r, Scheme 2 saves more power than Scheme 1. The reason is, in Scheme 2, the weak user adopts caching technique, resulting in high probability to eliminate the interference from strong user. This can enhance the system performance in terms of power consumption to some extent.

VI. Conclusion

This work studies the power minimization problem for the caching based downlink NOMA system with the consideration of each user's data rate constraint. To solve this non-convex optimization problem, we first transform the power requirements of users to that of BSs. With the power constraint transformation, the original power control problem can be transformed into a power minimization problem for BSs. The optimality and the uniqueness of the optimal solution for the newly formulated power control problem are analyzed. In addition, we show that the optimal objective function values of the two problems are the same. This motivates us to obtain the optimal solution to the original problem via solving the newly formulated problem. An iterative power control algorithm is thus designed. Simulation results show that cache-based NOMA under the proposed power control method saves more power than the optimal power-controlled NOMA system. In future, we will investigate the joint user clustering and power allocation for the downlink of caching based NOMA networks with an arbitrary number of cells and an arbitrary number of users per cell.

References

[1] V. W. S. Wong, R. Schober, D. W. K. Ng, and L. C. Wang, *Key technologies for 5G wireless systems*, 1st ed. Cambridge University Press, Apr. 2017.

[2] Z. Wei, J. Yuan, D. W. K. Ng, M. Elkashlan, and Z. Ding, "A survey of downlink non-orthogonal multiple access for 5G wireless communication networks," *ZTE Communications*, vol. 14, no. 4, pp. 17–25, Oct. 2016.

[3] Z. Q. Al-Abbasi and D. K. C. So, "Resource allocation in non-orthogonal and hybrid multiple access system with proportional rate constraint," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6309–6320, Oct. 2017.

[4] S. M. R. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, "Resource allocation for downlink NOMA systems: Key techniques and open issues," *https://arxiv.org/abs/1801.00121*, Dec. 2017.

[5] Y. Fu, Y. Chen, and C. W. Sung, "Distributed power control for the downlink of multi-cell NOMA systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6207–6220, Sep. 2017.

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

[7] W. Han, A. Liu, and V. K. N. Lau, "PHY-caching in 5G wireless networks: Design and analysis," *IEEE Communications Magazine*, vol. 54, no. 8, Aug. 2016.

[8] M. A. Tebbi and C. W. Sung, "Coded caching in partially cooperative D2D communication networks," *to appear in Proc. ICUMT*, pp. 148–153, Jan. 2018.

[9] J. A. Oviedo and H. R. Sadjadpour, "Leveraging edge caching in NOMA systems with QoS requirements," *https://arxiv.org/abs/1801.07430*, Jan. 2018.

[10] B. Perabathini, E. Bastug, M. Kountouris, M. Debbah, and A. Conte, "Caching at the edge: A green perspective for 5G networks," *in Proc. IEEE International Conference on Communication Workshop (ICCW)*, pp. 2830–2835, Jun. 2015.

[11] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*, 2nd ed. Athena Scientific, 2002.

[12] E. Seneta, *Non-negative matrices and Markov chains*, 2nd ed. Springer-Verlag, 1981.

[13] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

[14] ETSI, "3gpp tr 36.931 version 9.0.0 release 9," May. 2011.

[15] Y. Fu, L. Salaün, C. W. Sung, C. S. Chen, and M. Coupechoux, "Double iterative waterfilling for sum rate maximization in multicarrier NOMA systems," *IEEE International Conference on Communications (ICC)*, May. 2017.