Optimal user pairing in cache-based NOMA systems with index coding

Fu, Y., Shum, K., Sung, C. W. & Liu, Y.

Original citation & hyperlink:
Fu, Y, Shum, K, Sung, CW & Liu, Y 2019, Optimal user pairing in cache-based NOMA systems with index coding. in 2019 IEEE International Conference on Communications (ICC). IEEE, pp. 1-6, 2019 IEEE International Conference on Communications , Shanghai, China, 20/05/19.
https://dx.doi.org/10.1109/ICC.2019.8761754

DOI 10.1109/ICC.2019.8761754
ESSN 1938-1883

Publisher: IEEE

Optimal User Pairing in Cache-based NOMA Systems with Index Coding

Yaru Fu*, Kenneth W. Shum†, Chi Wan Sung‡, and Ye Liu§
*Information Systems Technology and Design, Singapore University of Technology and Design, Singapore
†School of Science and Engineering, The Chinese University of Hong Kong (Shenzhen), China
‡Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR
§Department of Engineering Science, University of Oxford, England, UK
Email: yaru_fu@sutd.edu.sg, wkshum@cuhk.edu.cn,
albert.sung@cityu.edu.hk, ye.liu@eng.ox.ac.uk

***Abstract*—The user pairing problem for cache-based time-slotted non-orthogonal multiple access (NOMA) system with index coding is investigated. During each time slot, the packets of two users are scheduled at the base station. In accordance with different cache information of the scheduled users, either superposition coding or index coding is applied for base station transmission. For some specific case, the superior performance on the aspect of power consumption of index coding compared to that of superposition coding is analyzed. Besides, the power saving of our design system when compared to that of the NOMA system with pure superposition coding is also demonstrated in a mathematical way. Subsequently, we show that the original user scheduling problem can be transformed in quadratic time into a minimum weight perfect matching problem of an undirected graph, which can be solved with time complexity $O(K^3)$, where K is the number of users. Based on this transformation, the feasibility of any given system is analyzed. Furthermore, we formulate the minimum weight perfect matching problem as an integer linear problem and solve it by integer linear programming. Numerical results validate the performance gains of our proposed system from the aspects of total transmit power and outage probability.**

***Index Terms*—Non-orthogonal multiple access (NOMA), caching, superposition coding, index coding, scheduling, minimum weight perfect matching.**

I. Introduction

As a promising candidate for the fifth generation (5G) wireless communication networks, non-orthogonal multiple access (NOMA) has attracted significant attention from both academia and industry. In contrast to conventional orthogonal multiple access (OMA) schemes, where each time-frequency resource block (RB) can only be occupied by a single user, NOMA provides a new horizon of thinking as it supports multiple users to use the same RB simultaneously. By so doing, NOMA is capable of having superior spectrum efficiency compared to OMA strategies [1].

Since 2013, NOMA has been studied from various perspectives. For example, [2] and [3] study the power minimization problem for single-carrier NOMA networks, in which distributed power allocation algorithm based on Yate's power control framework and multi-user non-cooperative game theory is adopted, respectively. Besides, the joint subcarrier and power allocation for sum-rate maximization problem in multi-carrier NOMA system are investigated by [4], [5]. Specifically, in [4], the rate maximization problem is proved to be NP-hard and solved via a Lagrangian duality assisted dynamic programming (LDDP) method. Meanwhile, a three-step resource allocation methodology with provable lower computation complexity than LDDP is designed in [5]. In addition, [6]–[8] investigate the user clustering, beamforming and power allocation problem for NOMA based multi-antenna systems. The work in [9] investigates the use of NOMA to support cell-edge users.

Recently, people are turning the research trend to cache-based NOMA systems in 5G wireless networks [10], [11]. It is demonstrated by [12] that cache-based NOMA obtains superior outage performance when compared with that of the conventional NOMA. In addition, the work in [13] shows that caching has the capability of improving system coverage and spectral efficiency. While [12], [13] focus on the single-cell case, in [14], the power minimization problem for a two-cell cache-based NOMA network is investigated, in which an iterative power control algorithm is proposed to solve the non-convex optimization problem optimally.

All of the aforementioned works [12]–[14] apply superposition coding at transmit side. To the best of our knowledge, this is the first work which considers cache-aided NOMA system with index coding [15], [16], in which the transmission mechanism of the base station is either superposition coding or index coding based on the cached information at users. The motivation of studying such an implementation is that, for some specific cases, the power requirement for successful decoding at the scheduled users with index coding is less than that with superposition coding.

To understand the potential of considering NOMA and index coding jointly, in this work, we study the user pairing problem for cache-aided NOMA network with the application of index coding, in which a time-slotted system is considered. During

This work was done when Yaru Fu worked in the Chinese University of Hong Kong and in City University of Hong Kong in 2018, and was supported in part by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China, under the project CityU 11216416. (Corresponding author: Chi Wan Sung)

each time slot, the desired packets of two users are scheduled at the base station. Based on different cache information of the scheduled users, either superposition coding or index coding is applied at base station. When two users request packets of their own from the base station while they cache the desired packet of each other, the superior performance on the part of power consumption of index coding compared with that of superposition coding is analyzed. Afterwards, we analyze the power saving of our proposed system when comparing with NOMA using pure superposition coding mathematically. Then, we show that the original user scheduling problem can be transformed into a minimum weight perfect matching problem of an undirected graph in quadratic time. Based on the formulated graph, the feasibility of any given system is analyzed. Finally, we formulate the minimum weight perfect matching problem as an integer linear problem, and solve it via integer linear programming. Numerical results indicate that our proposed cache-aided NOMA with index coding outperforms conventional cache-based NOMA and NOMA without caching in terms of total transmit power and outage probability. The basic model in this paper is subsequently extended in [17], where a suboptimal joint user clustering and power control algorithm is proposed with the consideration of inter-cluster interference.

The remaining parts of this article are outlined as follows. We present the system model and formulate the user scheduling problem in Section II, where we also prove the superiority of the designed system compared to the conventional NOMA network mathematically. In Section III, the original pairing problem is transformed into a minimum weight perfect matching problem in general graph, in which the feasibility of the system is analyzed. Moreover, we formulate the minimum weight perfect matching problem as an integer linear problem, which can be solved by integer linear programming. Numerical results are presented in Section IV. Section V concludes the main contributions of this work and predicts several possible research directions in future.

II. System Model

In this section, the system model of cache-based NOMA with index coding is first described, where the transmission mechanism at base station is specified. Then, we show the advantage of our designed system compared to conventional NOMA system in terms of power consumption mathematically. Afterwards, the user pairing problem formulation is presented.

A. System Description

Consider a single base station which wants to transmit K different packets to K users. Assume each packet contains a string of bits, and the size of the packets are the same. The K packets and K users are both indexed by $\mathcal{K} \triangleq \{1, 2, \dots, K\}$. Let s_k be the k-th packet and is the desired packet of user k, where $k \in \mathcal{K}$. Besides, we assume user k has stored in its cache a set of packets. Define $\mathcal{H}_k \subseteq \mathcal{K} \setminus \{k\}$ as the set of the indices of the cached packets at user k, e.g., given $\mathcal{H}_k = \{4, 5\}$, then user k has the information about packets s_4 and s_5 beforehand. In addition, each user is assumed to have a maximum cache capacity $\bar{C}$, meaning that each user is able to cache the packets of any $\bar{C}$ other users.

Let X be the signal[1] transmitted by the base station and denote by Y_k the received message of user k. Then

$$Y_k = \sqrt{g_k} X + N_k, \tag{1}$$

where g_k is the time-invariant link gain from the base station to user k and N_k is the additive Gaussian random noise signal with mean zero and variance η. The transmission of the base station is subject to a power constraint based on the third generation of partnership project (3GPP) configuration [18], so that the variance of X is no greater than P. Without loss of generality, we assume that the users are ordered in descending order of their link gains, i.e., $g_1 \geq g_2 \cdots \geq g_K$. For $k \in \mathcal{K}$, we define the channel gain of user k that normalized by noise power as $n_k \triangleq \eta / g_k$. We assume that the same modulation and coding schemes are applied for all users, so there is a common signal-to-interference-plus-noise ratio (SINR) threshold, Γ, that all packets need to meet for successful decoding. Specifically, Γ is in one-to-one correspondence of the transmission data rate constraint of each user, which is denoted by R. Based on Shannon capacity formula, we have

$$\Gamma = 2^{R/W} - 1, \tag{2}$$

in which W represents the system bandwidth.

We consider a time-slotted system and focus on the user pairing problem, in which the packets of two users are scheduled at base station in each time slot. Let $\mathcal{U}_n$ be the set of users that are scheduled in time slot n for transmission; therefore, we have $|\mathcal{U}_n| = 2$ for all n. Depending on the cache information, either superposition coding or index coding is used for transmission. For simplicity, we assume K is an even number. Therefore, $K/2$ time slots are needed. As a consequence, we have $n \in \{1, 2, \dots, K/2\}$.

In the following, we discuss the transmission mechanism at base station. We assume there is an encoder function named enc, which encodes the transmitted packet of user k, X_k, as follows:

$$X_k = \text{enc}(s_k), \tag{3}$$

in which $\text{enc}(s_k)$ is a random variable whose variance and mean are assumed to be unit and zero, respectively.

Assume user i and user j are scheduled within the same time slot. Therefore, for superposition coding, the transmitted signal at the base station is quoted below:

$$\begin{aligned} X &= \sqrt{p_i} \cdot \text{enc}(s_i) + \sqrt{p_j} \cdot \text{enc}(s_j) \\ &= \sqrt{p_i} X_i + \sqrt{p_j} X_j, \end{aligned} \tag{4}$$

[1]Note that X is the transmit signal, which can be constructed by either superposition coding or index coding. We will present the details later in this subsection.

where p_k is the transmit power of packet k where $k \in \mathcal{K}$. In accordance with (1), we obtain the received signals as follows:

$$\begin{aligned} Y_i &= \sqrt{g_i}(\sqrt{p_i}X_i + \sqrt{p_j}X_j) + N_i, \\ Y_j &= \sqrt{g_j}(\sqrt{p_i}X_i + \sqrt{p_j}X_j) + N_j. \end{aligned} \tag{5}$$

If index coding is applied, the transmitted signal at base station is expressed as follows:

$$X = \sqrt{p_{IC}} \cdot \mathrm{enc}(s_i \oplus s_j). \tag{6}$$

Here, $\oplus$ is binary addition, and p_{IC} is the transmit power.

Suppose a pair of users (i, j), where $i < j$, is scheduled for transmission in the same time slot. Besides, we define the binary vector $I_{ij} \triangleq (b_i, b_j)$ to indicate whether a user has the other user's packet in his cache. Specifically, $b_i = 1$ if and only if user i cached the packet of user j, and b_j is defined in the same way.

With above discussions, in the transmission scheme, the following mechanism is adopted: if $I_{ij} = (1, 1)$, we use index coding; otherwise, we use superposition coding. The reason is stated in the following proposition:

Proposition 1. *If $I_{ij} = (1, 1)$, then the power requirement for successful decoding with index coding is smaller than the power requirement for successful decoding with superposition decoding.*

Proof. We first consider the scenario where superposition coding is adopted. Since users i and j have each other's packet, no user suffers from inter-user interference. Therefore, at the receiver side, the following SINR conditions should be satisfied for successful decoding:

$$\frac{p_i}{n_i} \geq \Gamma, \text{and } \frac{p_j}{n_j} \geq \Gamma.$$

Obviously, the least required power under superposition coding is $p_{SC} = \Gamma(n_i + n_j)$.

Subsequently, we consider the case of index coding. At the base station side, the XOR operation is used among users i and j to form the coded packet. At the receive side, both users first decode this coded packet. Then, XOR is performed again at each user between the decoded packet and its cached packet to obtain his/her desired information. Although we need to ensure both users can decode the coded packet, only one constraint is needed, since $n_i \leq n_j$. The constraint is given below

$$\frac{p_{IC}}{n_j} \geq \Gamma, \tag{7}$$

where p_{IC} is the transmit power of the index coded packet. Therefore, the minimum required power for successful decoding with index coding in this time slot is $p_{IC} = n_j\Gamma$.

Since $p_{SC} - p_{IC} = n_i\Gamma > 0$, the proof is completed. □

In this work, we adopt the assumption that any packet that are overheard is not used in decoding, which means the decoder of user k only uses the received signal in time slot n if $k \in \mathcal{U}_n$. With aforementioned definitions and analysis, we specifically distinguish the following cases:

1) $I_{i,j} = (0, 0)$ or $(1, 0)$. In this scenario, the base station applies superposition coding. In addition, user i adopts successive interference cancellation (SIC). For successful decoding at both users, the following constraints have to be satisfied:

$$\begin{aligned} \frac{p_i}{n_i} &\geq \Gamma, \\ \frac{p_j}{p_i + n_j} &\geq \Gamma, \end{aligned} \tag{8}$$

where p_i and p_j are the assigned power for transmission of packets i and j, respectively. Based on (8), the least required transmit power in this time slot can be obtained as $(n_i + n_j + n_i\Gamma)\Gamma$.

2) $I_{ij} = (0, 1)$. Similarly, superposition coding is used in this case. Since user j cached the packet of user i, which indicates no matter what the received SINR is, user j can subtract packet i from his received message. Therefore, we have the following constraint for user j:

$$\frac{p_j}{n_j} \geq \Gamma. \tag{9}$$

At user i, SIC is applied to decode its intended signal. In detail, user i first decodes the signal of user j, subtracting this part and then decoding its own signal. Therefore, the following constraints have to meet:

$$\frac{p_j}{p_i + n_i} \geq \Gamma, \tag{10}$$

$$\frac{p_i}{n_i} \geq \Gamma. \tag{11}$$

Solving the above inequalities, the minimum total power required in this time slot is then given by $n_i\Gamma + \max\{n_j\Gamma, n_i\Gamma(1 + \Gamma)\}$. Note that (10) is always true in conventional NOMA. However, it may not hold for NOMA system with caching due to the fact that the weak user stores the message of the strong user.

3) $I_{ij} = (1, 1)$. In this case, index coding is used. According to Proposition 1, the minimum required power in this time slot is $n_j\Gamma$.

With aforementioned analysis, we give the definition of feasible schedule in each time slot, which is given as follows:

Definition 2. *The scheduling in a time slot is feasible if and only if the minimum required power in this time slot is no great than the power constraint of base station, i.e., P.*

B. Cache-aided NOMA with Index Coding versus Conventional NOMA

In this subsection, we show the superior performance of above discussed cache-aided NOMA with index coding when compared to conventional NOMA where $\mathcal{H}_k = \varnothing$ for $k \in \mathcal{K}$ in terms of power consumption.

It is easy to see that the conventional NOMA is equivalent to the case where $I_{i,j} = (0, 0)$, whose total required power for a pair of users is given by $(n_i + n_j + n_i\Gamma)\Gamma$. Define p_s as the saved power of cache assisted NOMA compared to conventional NOMA under the same system configurations.

Therefore, the power saving of cache-aided NOMA with index coding can be summarized as follows:

1) when $I_{i,j} = (1, 0)$, $p_s = 0$, which means in this case, cache based NOMA with index coding requires the same power as that of conventional NOMA.
2) when $I_{i,j} = (0, 1)$, $p_s = \min\{n_i\Gamma^2, (n_j - n_i)\Gamma\}$.
3) when $I_{ij} = (1, 1)$, $p_s = (1 + \Gamma)n_i\Gamma$.

C. Problem Formulation

Before introducing the problem to be solved, we specify the definition of feasible system as follows:

Definition 3. *Given a realization of link gains, we say that the system is feasible if and only if there exists $K/2$ feasible schedules as defined in Definition 2 such that each user $k \in \mathcal{K}$ gets its intended packet.*

In this paper, we aim to minimize the sum power consumption of the $K/2$ time slots as mentioned in Definition 3 to ensure that every user obtains its required packet.

III. Minimum Weight Perfect Matching in a Graph

In this section, we first show that the original user scheduling problem can be transformed into a minimum weight perfect matching problem in an undirected graph. Based on the constructed graph, the feasibility of any given system is analyzed. Furthermore, we formulate the minimum weight perfect matching problem as an integer linear problem, which can be solved by using integer linear programming.

A. Construction of Graph Problem

The user scheduling problem can be transformed to a graph problem in quadratic time. Consider an undirected graph with K vertices, which correspond to the K distinct users. We examine every possible pair of users, and compute the minimum required power when they are scheduled within the same time slot. Consider any user pair (i, j), and define their required power as p_{ij}. According to Definition 2, there is an edge connecting vertices i and j if and only if $p_{ij} \le P$. Each edge that satisfies $p_{ij} \le P$ in the graph is assigned by a non-negative weight of p_{ij}. Since there are $\binom{K}{2}$ edges, the construction of the undirected weighted graph has a time complexity of $O(K^2)$. Let E be the set of all feasible edges. A subset $E' \subseteq E$ is called a *matching* if each vertex $k \in \mathcal{K}$ has at most one incident edge in E'. Furthermore, a matching E' is said to be *perfect* if each node $k \in \mathcal{K}$ has exactly one incident edge in E'.

According to the definition of perfect matching, if a given problem instance is feasible, then there exist a perfect matching in the graph. In that case, our objective is to find a perfect matching that has the minimum sum weight.

B. Feasibility Analysis

Given any problem instance, let $G = (\mathcal{K}, E, p)$ be the constructed undirected weighted graph according to Section III-A, where p represents the edge cost, which represents the power requirement for the corresponding pair of users. Define T as the $K \times K$ Tutte matrix of graph G, and the element in its i-th row and j-th column is given as follows:

$$T(i,j) = \begin{cases} 0 & \text{if } (i,j) \notin E \\ 1 & \text{if } (i,j) \in E,\ i < j \\ -1 & \text{if } (i,j) \in E,\ i > j \end{cases}. \tag{12}$$

The feasibility of any given system can then be characterized by the following proposition:

Proposition 4. *The system is feasible if and only if the determinant of the Tutte matrix T is non-zero, i.e.,* $\det(T) \neq 0$.

Proof. Based on [19], for any given G with even $|\mathcal{K}|$, the determinant of the Tutte matrix T is non-zero if and only if G contains a perfect matching. According to Section III-A, if a perfect matching exists, the system is feasible. □

C. Integer Linear Problem Formulation

Suppose a perfect matching in G exists. Then, our goal is to find a perfect matching E^* that has minimum weight $p(E^*)$. Incident vector $x \in \{0, 1\}^E$ is used in this work to represent matching $E' \subseteq E$. For any $S \subseteq \mathcal{K}$, let $f(S) = \{(i, j) \in E | i \in S, j \in \mathcal{K} - S\}$, which can be regarded as the set of boundary edges of S. Note that for single node, we assume $f(k) = f(\{k\})$, where $k \in \mathcal{K}$. Based on the definition of perfect matching and the aforementioned analysis, we formulate the minimum weight perfect matching as the following integer linear problem:

$$\min \sum_{e \in E} p_e x_e \tag{13}$$

subject to

$$\begin{aligned} &C1: \ x(f(k)) = 1, \ \forall\ k \in \mathcal{K}, \\ &C2: \ x_e \in \{0, 1\}, \ \forall\ e \in E, \end{aligned} \tag{14}$$

where $C1$ indicates that each node has exactly one incident edge, $C2$ shows that the indicator variable of edge e, x_e, is a binary variable. The minimization problem (13) can be solved by integer linear programming [20]. Alternatively, it can also be solved in polynomial time by the classic Edmond's algorithm [21], whose time complexity can be further reduced to $O(K^3)$ [22].

IV. Simulation Results

In this section, we use Monte-Carlo simulation to compare the performance of our proposed cache-based NOMA system with index coding to that of the conventional cache based NOMA and NOMA without caching systems. The cell radius R is assumed to be 500 meters. Within the cell, there is one base station fixed at the cell center and serves K uniformly distributed users. We set the system bandwidth to be $W = 5$ MHz. In addition, the noise power density is assumed to be -174 dBm/Hz. Besides, we assume that the minimum data rate of each user is 10 Mbits/s. For the radio propagation model, both the large scale fading and the small scale fading are taken into account. Specifically, the large-scale fading is assumed to be distance-dependent and it is

TABLE I
SIMULATION PARAMETERS

Parameters	Value
Cell radius	500 m
Distance-dependent path loss	$128.1 + 37.6 \log_{10} d$ dB, d is in km
Small-scale fading	Rayleigh fading with variance 1
The distribution of users	Randomly uniform distribution
Power constraint per transmission, P	2 W
Noise power density	-174 dBm/Hz
Total power budget, P_{total}	2 W
System bandwidth, W	5 MHz
Number of users, K	4 to 12
Data rate calculation	Shannon's capacity formula
Successful decoding threshold, Γ	3
Cache capacity per user, $\overline{C}$	2 to 6

given by $128.1+37.6\log_{10} d$, where d is in kilometers, which represents the distance between the base station and the user. For the small-scale fading, independent Rayleigh fading with variance 1 is applied among users. In addition, we consider a total power budget P_{total} for the $K/2$ time slots, which is set to be the same as the power constraint of base station, P. We summarize the simulation parameters in Table I.

A. Power Consumption

In this subsection, we compare the power consumption of our proposed system to that of the cache-based NOMA and conventional NOMA system. Specifically, in the cache-based NOMA, the transmission mechanism at the base station is always superposition coding. Meanwhile, in conventional NOMA, caching is not applicable, which means the cache capacity of conventional NOMA is $\bar{C} = 0$. To obtain the total power consumption for aforementioned three systems, we randomly generate 100,000 link gain instances. Each data is calculated by averaging the power consumptions of the instances where each of the systems requires a transmit power that is no great than P_{total}. Besides, we define "Power Saving 1" as the power saving ratio of our cache based NOMA with index coding when compared to that of conventional NOMA. The calculating method is quoted below

$$\frac{P_{\text{total}}(\text{NOMA}) - P_{\text{total}}(\text{C-NOMA-I})}{P_{\text{total}}(\text{NOMA})} \times 100\%, \qquad (15)$$

where $P_{\text{total}}(\text{NOMA})$ and $P_{\text{total}}(\text{C-NOMA-I})$ represents the average required transmit power for successful decoding of all users' packets of conventional NOMA and our designed NOMA, respectively. Similarly, we define "Power Saving 2" as the power saving ratio of cache based NOMA with index coding when compared to that of conventional cache-based NOMA.

Fig. 1 shows the total transmit power of the above mentioned three systems, where the cache capacity $\bar{C}$ is set to be 2, which means that each user could cache the packets of any two of the other users. From Fig. 1, it is easy to see that, with more users, the required transmit power of all systems increase. Besides, for any given K, the required power for successful decoding of cache based NOMA with index coding is less than the needed power of both NOMA without caching and conventional cache-based NOMA network. In addition, when the number of users is smaller, the power saving ratio of the cache based NOMA with index coding is larger. For example, when $K = 4$, cache based NOMA with index coding saves power by 25.59% and 10.42% comparing to that of NOMA without caching and conventional cache-based NOMA systems. The reason is when K is smaller, the probability for weak users to cache the packets of strong users and the ratio of the case where $I_{i,j} = (1,1)$ are higher.

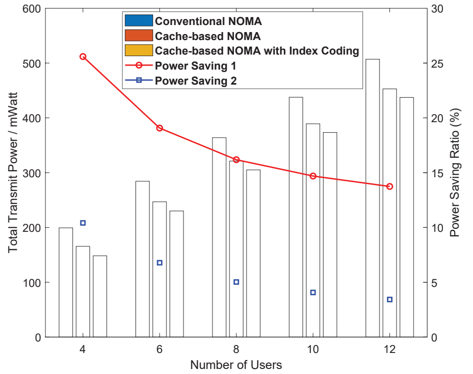

Fig. 1. Power consumption v.s. The number of users, K

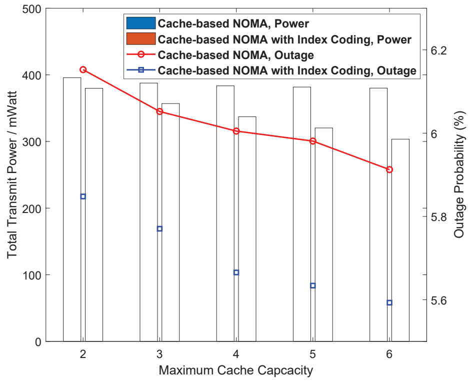

Fig. 2. Power consumption / Outage probability v.s. The cache capacity of users, $\bar{C}$

In addition, we also demonstrate the relationship between the power consumption of two cache-based NOMA systems and the maximum cache capacity of each user $\bar{C}$. This is shown by Fig. 2 (left side), in which the number of users is assumed to be $K = 10$. The x-axis represents the cache capacity of each user, while the left y-axis indicates the total power consumption. According to Fig. 2, we see that, given any $\bar{C}$, our designed cache based NOMA with index coding requires less power than that of conventional cache based NOMA. Besides, the required power of both systems decreases with the increasing of $\bar{C}$. Moreover, the larger the value of $\bar{C}$, the less total required power of our designed NOMA with index coding scheme. For example, when $\bar{C} = 6$, cache based NOMA with index coding saves power by 20.09% when compared to that of the case where $\bar{C} = 2$.

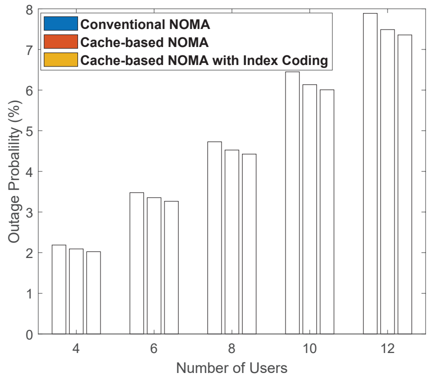

Fig. 3. Outage probability v.s. The number of users, K

B. Outage Probability

In this subsection, we compare the outage performance of our designed NOMA system to conventional NOMA and cache-based NOMA. As aforementioned, we randomly generate 100,000 instances. For each system, the instances that require more transmit power than P_{total} are called outage instances. The outage probability of each system is equal to the number of outage instances divide the total number of instances.

Fig. 3 plots the outage probability of three systems against the number of users. Obviously, the outage ratios of all systems increase with the increasing of users' number. In addition, for any given K, the outage probability of the cache based NOMA with index coding is less than that of the other two schemes, since our designed scheme requires the least total transmit power.

The relationship between outage performance of two cache based NOMA schemes and the cache capacity of user is shown in Fig. 2 (right side), from which we conclude that with the increasing of user's cache capacity $\bar{C}$, the outage probability of both two cache based systems are decreasing. Besides, for any given $\bar{C}$, the outage performance of our designed NOMA scheme is better than that of conventional cache-based NOMA.

V. Conclusion

In this work, we studied the optimal user scheduling problem for cache based NOMA system with index coding. A time-slotted system was considered. During each time slot, the packets of two users were scheduled at the base station. In accordance with different cache information of the scheduled users, either superposition coding or index coding was applied at the base station. We demonstrated the power efficiency of index coding. Furthermore, we analyzed the power saving of our design system compared to that of conventional NOMA mathematically. Afterwards, we showed that the original user pairing problem could be transformed into a minimum weight perfect matching problem of a general graph. Based on that, the feasibility of any given system was analyzed. Moreover, we formulated the minimum weight perfect matching problem as an integer linear problem, and solved it by integer linear programming. Simulation results validate the potential gains of applying index coding in cache-based NOMA system. In future, the cache decision optimization at the base station as well as user recommendation will be considered in cache assisted NOMA networks.

References

[1] 3rd Generation Partnership Project (3GPP), "Study on downlink multiuser superposition transmission for LTE," Mar. 2015.

[2] 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.

[3] C. W. Sung and Y. Fu, "A game-theoretic analysis of uplink power control for a non-orthogonal multiple access system with two interfering cells," *IEEE Vehicular Technology Conference (VTC)*, pp. 1–5, May 2016.

[4] L. Lei, D. Yuan, C. K. Ho, and S. Sun, "Power and channel allocation for non-orthogonal multiple access in 5G systems: tractability and computation," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8580–8594, Dec. 2016.

[5] Y. Fu, L. Salaün, C. W. Sung, and C. S. Chen, "Subcarrier and power allocation for the downlink of multicarrier NOMA systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 11 833–11 847, Dec. 2018.

[6] H. Wang, R. Zhang, R. Song, and S. Leung, "A novel power minimization precoding scheme for MIMO-NOMA uplink systems," *IEEE Communications Letters*, vol. 22, no. 5, pp. 1106–1109, May 2018.

[7] Y. Liu, H. Xing, C. Pan, A. Nallanathan, M. Elkashlan, and L. Hanzo, "Multiple antenna assisted non-orthogonal multiple access," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 17–23, Apr. 2018.

[8] Y. Yuan and Z. Ding, "Outage constrained secrecy rate maximization design with SWIPT in MIMO-CR systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 5475–5480, Jun. 2018.

[9] Q. Guo, C. W. Sung, Y. Chen, and C. S. Chen, "Power control for coordinated NOMA downlink with cell-edge users," *IEEE WCNC*, pp. 1–6, Apr. 2018.

[10] Z. Ding, P. Fan, G. K. Karagiannidis, R. Schober, and H. V. Poor, "On the application of NOMA to wireless caching," *IEEE International Conference on Communications (ICC)*, pp. 1–7, May 2018.

[11] L. Xiang, D. W. K. Ng, X. Ge, Z. Ding, V. W. Wong, and R. Schober, "Cache-aided non-orthogonal multiple access," *IEEE International Conference on Communications (ICC)*, pp. 1–7, May 2018.

[12] Z. Zhao, M. Xu, W. Xie, Z. Ding, and G. K. Karagiannidis, "Coverage performance of NOMA in wireless caching networks," *IEEE Commun. Lett.*, vol. 27, no. 7, pp. 1458–1461, Jul. 2018.

[13] Z. Ding, P. Fan, G. K. Karagiannidis, R. Schober, and H. V. Poor, "NOMA assisted wireless caching: Strategies and performance analysis," *IEEE Transactions on Communications*, vol. 66, no. 10, pp. 4854–4876, Oct. 2018.

[14] Y. Fu, H. Wang, and C. W. Sung, "Optimal power allocation for the downlink of cache-aided NOMA systems," *IEEE Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, Oct. 2018.

[15] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1479–1494, Mar. 2011.

[16] M. Dai, K. W. Shum, and C. W. Sung, "Data dissemination with side information and feedback," *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4708–4720, Sep. 2014.

[17] Y. Fu, Y. Liu, H. Wang, Z. Shi, and Y. Liu, "Mode selection between index coding and superposition coding in cache-based NOMA networks," *to appear in IEEE Communications Letters*, 2019.

[18] 3rd Generation Partnership Project (3GPP), "TR 25.951 Release 11: Technical specification group radio access network," Sep. 2012.

[19] A. Abboud, "Perfect matching," *Matchings in graphs-Stanford CS Theory*, Apr. 2014.

[20] W. T. Tutte, *Graph Theory*, 2nd ed. Cambridge University Press, 2001.

[21] J. Edmonds, "Paths, trees and flowers," *Canadian Journal of Mathematics*, vol. 17, pp. 449–467, 1965.

[22] E. Lawler, *Combinatorial Optimization: Networks and Matroids*. Dover Publications, 2011.