Design of Placement Delivery Arrays for Coded Caching With Small Subpacketizations and Flexible Memory Sizes

Xianzhang Wu, Minquan Cheng, Congduan Li, *Member, IEEE*, and Li Chen, *Senior Member, IEEE*

***Abstract*—Coded caching is an emerging technique to reduce the data transmission load during the peak-traffic times. In such a scheme, each file in the data center or library is divided into a number of packets to pursue a low broadcasting rate based on the designed placements at each user's cache. However, the implementation complexity of this scheme increases with the number of packets. It is crucial to design a scheme with a small subpacketization level, while maintaining a relatively low transmission rate. Recently, a combinatorial structure called placement delivery array (PDA) was proposed as an effective tool to design coded caching schemes with a relatively low subpacketization level. This paper proposes a novel PDA construction by selecting proper orthogonal arrays (POAs), which generalizes the existing construction but with a more flexible memory size. Based on the proposed PDA construction, an effective transform is further proposed to enable a coded caching scheme to achieve a smaller subpacketization level. Moreover, two new coded caching schemes with the coded placement are derived. It is shown that the proposed schemes can yield a lower subpacketization level or transmission rate over the benchmark schemes.**

***Index Terms*—Coded caching, placement delivery array, proper orthogonal array, subpacketization level.**

I. Introduction

THE dramatically increasing demands of video streaming services generate challenges to the central servers for ensuring a smooth data transmission, especially during the peak hours. Coded caching system has been proposed as a promising technology to reduce the data transmission load during the peak hours by utilizing ample cache memories that are available at the users. In the centralized coded caching system, a central server containing N files of the same size is connected to K users over an error free broadcasting channel. Each user is assumed to have a cache memory with a size of M files, where $M < N$. It consists of two phases: the placement phase which occurs during the off-peak hours, and the delivery phase which occurs during the peak hours. In the placement phase, the server sends the properly designed contents to the cache of each user without knowledge of their demands. In the delivery phase, each user requests an arbitrary file, and the server will then broadcast some coded packets to them so that each user can recover its desired file with the assistance of the contents in its own cache. The worst case minimal broadcasting load over all possible demands is defined as the *transmission rate* (or *rate*) R, i.e., the least number of files that must be communicated so that any demands can be satisfied. A coded caching scheme is called an F-division scheme if each file can be equally divided into F packets, which is called the *subpacketization level*. If the packets are cached directly without coding in the placement phase, it is called an *uncoded placement*. Otherwise, it is called a *coded placement*. We summarize the prior work and the state of the arts as follows.

A. Prior Work and the State of the Arts

The information theoretic coded caching model proposed by Maddah-Ali and Niesen [1] is realized by a combinatorial design in the placement phase and a linear coding in the delivery phase, which is referred to as the MN scheme. It was shown that the transmission rate achieved by this scheme is within a factor of 12 of a cut-set-type lower bound [1], which does not assume any restrictions on the placement or the delivery phase. Following the seminal work of [1], there have been some improvements in terms of both the achievable rates [2], [3] and the lower bounds [4], [5], where the gap has been shorten significantly. In fact, under the constraints of uncoded placement, the transmission rate of the MN scheme was shown to be exactly optimal when $K \leq N$ in [6] and [7]. This is realized by the use of the well-studied index coding problems [8], [9]. There are some other discussion focusing

Manuscript received 4 January 2022; revised 10 May 2022 and 28 August 2022; accepted 12 September 2022. Date of publication 19 September 2022; date of current version 18 November 2022. This work is sponsored by the National Natural Science Foundation of China (NSFC) with project IDs 62071498, 61901534, 62271514, U21A20474 and 62061004, the Science, Technology and Innovation Commission of Shenzhen Municipality with project IDs JCYJ20190807155617099 and JCYJ20210324120002007, the Guangdong Basic and Applied Basic Research Foundation with project ID 2019B1515120032, and the Guangxi Natural Science Foundation with project ID DA035087. An earlier version of this paper was presented in part at the IEEE International Symposium on Information Theory (ISIT), 2022 [DOI: 10.1109/ISIT50566.2022.9834597]. The associate editor coordinating the review of this article and approving it for publication was M. Ji. *(Corresponding authors: Li Chen; Congduan Li.)*

Xianzhang Wu and Congduan Li are with the School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen 518107, China (e-mail: wuxzh7@mail2.sysu.edu.cn; licongd@mail.sysu.edu.cn).

Minquan Cheng is with the Guangxi Key Laboratory of Multi-Source Information Mining and Security, Guangxi Normal University, Guilin 541004, China (e-mail: chengqinshi@hotmail.com).

Li Chen is with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China (e-mail: chenli55@mail.sysu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCOMM.2022.3207838.

Digital Object Identifier 10.1109/TCOMM.2022.3207838

TABLE I
SUMMARY OF SOME KNOWN SCHEMES

Schemes and Parameters	User Number K	Caching Ratio $\frac{M}{N}$	Rate R	Subpacketization Level F
MN scheme in [1], any k, $t \in \mathbb{N}^+$ with $t < k$	k	$\frac{t}{k}$	$\frac{k-t}{1+t}$	$\binom{k}{t}$
Scheme in [14], any k, $t, c \in \mathbb{N}^+$ with $t < k$	ck	$\frac{t}{k}$	$\frac{c(k-t)}{1+t}$	$\binom{k}{t}$
Scheme in [24], any k, $t \in \mathbb{N}^+$ with $t < k$	$\binom{k}{t+1}$	$1 - \frac{t+1}{\binom{k}{t}}$	$\frac{k}{\binom{k}{t}}$	$\binom{k}{t}$
	$\binom{k}{t}$	$\frac{t}{k}$	$\frac{\binom{k}{t+1}}{k}$	k
Scheme in [22], any $a, b, m, \lambda \in \mathbb{N}^+$ with $a < m$,$b < m$, and $\lambda < \min\{a, b\}$	$\binom{m}{a}$	$1 - \frac{\binom{a}{\lambda}\binom{m-a}{b-\lambda}}{\binom{m}{b}}$	$\frac{\binom{m}{a+b-2\lambda}\binom{a+b-2\lambda}{a-\lambda}}{\binom{m}{b}}$	$\binom{m}{b}$
Scheme in [16], any m, t, $k \in \mathbb{N}^+$, prime power q with $m + t \leq k$	$\begin{bmatrix} k-t+1 \\ 1 \end{bmatrix}_q$	$1 - \frac{q^{m+1}\begin{bmatrix} k-t \\ m+1 \end{bmatrix}_q}{\begin{bmatrix} k-t+1 \\ m+1 \end{bmatrix}_q}$	$\frac{q^{m+1}\begin{bmatrix} k-m-t \\ 1 \end{bmatrix}_q}{m+2}$	$\frac{\begin{bmatrix} k-t+1 \\ m+1 \end{bmatrix}_q \prod_{i=0}^{m}(q^{m+1}-q^i)}{(q-1)^{m+1}(m+1)!}$
Scheme in [17], any m, t, $k \in \mathbb{N}^+$, prime power q with $m + t \leq k$	$\begin{bmatrix} k \\ t \end{bmatrix}_q$	$1 - \frac{\begin{bmatrix} k-t \\ m \end{bmatrix}_q}{\begin{bmatrix} k \\ m+t \end{bmatrix}_q}$	$\frac{\begin{bmatrix} k \\ m \end{bmatrix}_q}{\begin{bmatrix} k \\ m+t \end{bmatrix}_q}$	$\begin{bmatrix} k \\ m+t \end{bmatrix}_q$
Scheme in [18], any m, q, $t \in \mathbb{N}^+$ with $t < m$ and $q \geq 2$	$\binom{m}{t}q^t$	$1 - (\frac{q-1}{q})^t$	$(q-1)^t$	q^m
	$\binom{m}{t}q^t$	$1 - \frac{1}{q^t}$	$\frac{1}{(q-1)^t}$	$(q-1)^t q^m$
Scheme in [25], any m, t, z, $q \in \mathbb{N}^+$ with $t < m$, $z < q$ and $q \geq 2$	$\binom{m}{t}q^t$	$1 - (\frac{q-z}{q})^t$	$\frac{(q-z)^t}{\lfloor \frac{q-1}{q-z} \rfloor^t}$	$\lfloor \frac{q-1}{q-z} \rfloor^t q^m$
Scheme in [26], any m, t and $q \in \mathbb{N}^+$ with $t < m$ and $q \geq 2$	$\binom{m}{t}q^t$	$1 - (\frac{q-1}{q})^t$	$(q-1)^t$	q^{m-1}

on the transmission rate. For example, Gómez-Vilardebó [10] proposed a scheme that achieves the best rate-memory region under the constraints of $N \leq K \leq \frac{N^2+1}{2}$ and coded placement. Based on interference elimination and coded placement strategy, Tian and Chen [11] proposed a scheme that yields an improved transmission rate with $K \geq N$. To reduce the average transmission rate in practical regimes of finite subpacketization level, Ji it *et al.* [12] proposed some decentralized coded caching schemes based on hierarchical greedy local graph coloring. Jin *et al*. [13] considered a class of coded caching schemes specified by a file partition parameter, which is further optimized to minimize the average transmission rate with an arbitrary file popularity.

It is well known that the transmission rate and the subpacketization level are often a tradeoff in the design of coded caching schemes [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30]. To the best of our knowledge, most existing schemes aiming to reduce the subpacketization level are constructed under uncoded placement. With a large number of users, the tradeoff between the subpacketization level and the transmission rate was first derived by Shanmugam *et al*. in [14]. However, for an arbitrary number of users, it is still difficult to characterize the tradeoff between them. It was shown in [24] that for the fixed transmission rate (the rate of the MN scheme), the MN scheme has a minimum subpacketization level of $F = \binom{K}{\frac{KM}{N}}$. It can be seen that its subpacketization level increases exponentially with the number of users, which makes it impractical for large networks. Therefore, it is important to reduce the subpacketization level of the MN scheme, while maintaining a relatively low transmission rate. In particular, Yan *et al*. [21] proposed a combinatorial structure called the placement delivery array (PDA) to address the subpacketization bottleneck of coded caching, and constructed two new classes of coded caching schemes with a reduced subpacketization level over that of the MN scheme. Shangguan *et al*. [18] later showed that many previously existing coded caching schemes could also be represented by the PDA. With the introduction of PDA, various coded caching schemes with a lower subpacketzation level than the MN scheme were proposed in [24], [25], [26], [27], [28], [29], and [30]. Some other combinatorial construction methods for reducing the subpacketization level were realized by the use of projective geometry [17], Ruzsa-Szemerédi graphs [19], hypergraphs [18], linear block codes [23], strong edge coloring of bipartite graphs [22], and combinatorial design theory [15]. Recently, Cheng it *et al.* [25] generalized the construction of [18] and obtained a coded caching scheme with a flexible memory size. Based on the construction of [18], they further proposed some new coded caching schemes with a reduced subpacketization level by using orthogonal arrays [26]. Both of them achieve a good performance in either the subpacketization level or the transmission rate. Some of the above mentioned schemes are summarized in Table I, where $\begin{bmatrix} k \\ t \end{bmatrix}_q = \frac{(q^k-1)\cdots(q^{k-t+1}-1)}{(q^t-1)\cdots(q-1)}$ for any positive integers k, t and a prime power q.

B. Contribution and Organization of This Paper

This paper considers the construction of PDAs for the centralized coded caching scheme under the constraint of $K \leq N$, aiming to achieve a low transmission rate with a

TABLE II
NEW PDA SCHEMES

Schemes and Parameters	User Number K	Caching Ratio $\frac{M}{N}$	Rate R	Subpacketization Level F
New scheme in *Theorem 1*, any $q, z, m, t \in \mathbb{N}^+$ with $q \geq 2$, $z < q$ and $t < m$	$\binom{m}{t} q^t$	$1 - (\frac{q-z}{q})^t$	$\frac{(q-z)^t}{\lfloor \frac{q-1}{q-z} \rfloor^t}$	$\lfloor \frac{q-1}{q-z} \rfloor^t q^{m-1}$
New scheme in *Theorem 2*, any $q, z, m, t \in \mathbb{N}^+$ with $q \geq 2$, $z < q$ and $t < m$	$[\binom{m-1}{t} \lfloor \frac{q-1}{q-z} \rfloor^t + \binom{m}{t} - \binom{m-1}{t}] q^t$	$1 - (\frac{q-z}{q})^t$	$(q-z)^t$	q^{m-1}
New scheme in *Theorem 3*, any $q, m, t \in \mathbb{N}^+$ with $q \geq 2$ and $t < m$. z_r^* denotes the minimal integer in the set $\mathcal{G}_r = \{z \mid \lfloor \frac{q-1}{q-z} \rfloor = r, z \in [1 : q-1]\}$, where $r \in [1 : q-1]$	$\binom{m}{t} q^t$	$\frac{1-(\frac{q-z_r^*}{q})^t}{1-(\frac{q-z_r^*}{q})^t + (\frac{q-z}{q})^t}$	$\frac{(q-z)^t}{\lfloor \frac{q-1}{q-z} \rfloor^t [1-(\frac{q-z_r^*}{q})^t + (\frac{q-z}{q})^t]}$	$\lfloor \frac{q-1}{q-z} \rfloor^t q^{m-1} [1 + (\frac{q-z}{q})^t - (\frac{q-z_r^*}{q})^t]$
New scheme in *Theorem 4*, any $q, m, t \in \mathbb{N}^+$ with $q \geq 2$ and $t < m$. z_r^* denotes the minimal integer in the set $\mathcal{G}_r = \{z \mid \lfloor \frac{q-1}{q-z} \rfloor = r, z \in [1 : q-1]\}$, where $r \in [1 : q-1]$	$[\binom{m-1}{t} \lfloor \frac{q-1}{q-z} \rfloor^t + \binom{m}{t} - \binom{m-1}{t}] q^t$	$\frac{1-(\frac{q-z_r^*}{q})^t}{1-(\frac{q-z_r^*}{q})^t + (\frac{q-z}{q})^t}$	$\frac{(q-z)^t}{1-(\frac{q-z_r^*}{q})^t + (\frac{q-z}{q})^t}$	$q^{m-1}[1 + (\frac{q-z}{q})^t - (\frac{q-z_r^*}{q})^t]$

smaller subpacketization level and a more flexible cache size. The technical contributions are in three folds.

First, by defining the proper orthogonal array (POA), we propose a novel construction of PDAs, based on which, the corresponding coded caching scheme (as stated in *Theorem 1*) can reduce the number of packets by a factor of q without sacrificing the transmission rate of the existing scheme of [25]. It also yields a more flexible memory size over the scheme of [26]. Table II presents the advanced features of the proposed PDA schemes, where the conclusion of *Theorem 1* can be seen. The proposed construction can be seen as a generalization of the work of [26], but it requires a delicate selection of the POAs.

Secondly, elaborating the use of POAs, we present an effective transform which leads to another coded caching scheme (as stated in *Theorem 2* and also summarized in Table II). By appropriately designing the baseline arrays to satisfy the PDA constraints, we show that when the cache memory ratio is greater than $\frac{1}{2}$, the proposed scheme can work for a larger number of users and a smaller subpacketization level compared with the scheme in *Theorem 1*.

Finally, since in the proposed PDAs some packets have no multicasting opportunities in the delivery phase, the uncoded placement is further modified into the coded placement. Consequently, two coded caching schemes (as stated in *Theorems 3* and *4*, and summarized in Table II) with a smaller subpacketization level are further proposed.

The rest of the paper is organized as follows. Section II formulates the coded caching problem and reviews some related results. The new PDA schemes are presented in Section III. Performance analyses of the proposed PDA schemes are given in Section IV. Finally, Section V concludes the paper.

II. SYSTEM MODEL AND RELATED RESULTS

This section briefly reviews the centralized coded caching system and some existing PDA constructions via OAs. First, some key notations are introduced as follows.

Notations: Let bolded capital letters, bolded lower-case letters, and curlicue letters denote arrays, vectors, and sets, respectively. Symbol $\oplus$ represents the exclusive-or (XOR) operation. Let $\mathbb{N}^+$ denote the set of positive integers. The sets of consecutive integers are denoted as $[x : y] = \{x, x+1, \ldots, y\}$. We use $\binom{[0:m-1]}{t}$ to represent the collection of all subsets of $[0 : m-1]$ with size t, i.e., $\binom{[0:m-1]}{t} = \{\mathcal{S} \mid \mathcal{S} \subseteq [0 : m-1], |\mathcal{S}| = t\}$. Given an $l \times m$ matrix $\mathbf{F}$ and a subset $\mathcal{S} \subseteq [0 : m-1]$, let $\mathbf{F}|_{\mathcal{S}}$ denote a submatrix of $\mathbf{F}$, which is obtained by taking all the columns indexed by $j \in \mathcal{S}$. Let $\mathbf{P}(i, j)$ denote the entry of array $\mathbf{P}$ with row and column indexed by i and j, respectively. Further let $(\mathbf{A}_0; \mathbf{A}_1; \ldots; \mathbf{A}_n)$ denote an array obtained by arranging arrays (or row vectors) $\mathbf{A}_0, \mathbf{A}_1, \ldots, \mathbf{A}_n$ from top to bottom. E.g., $(\mathbf{A}_0; \mathbf{A}_1) = \begin{pmatrix} \mathbf{A}_0 \\ \mathbf{A}_1 \end{pmatrix}$. Finally, all the vectors in examples are written as strings. E.g., (1, 0, 1, 0) is written as (1010).

A. Centralized Coded Caching System

In a centralized coded caching system, a server containing N files of equal size is connected to K users through an error free shared link, as shown in Fig.1. Each user has a cache of size M files, where $M < N$. The N files and K users are denoted by $\mathcal{W} = \{W_0, W_1, \ldots, W_{N-1}\}$ and $\mathcal{K} = [0 : K-1]$, respectively. An F-division (K, M, N) coded caching scheme consists of two phases as follows.

• **Placement Phase**: Each file is divided into F equal packets, i.e., $W_n = \{W_{n,j} | j \in [0 : F-1]\}$, $n \in [0 : N-1]$.

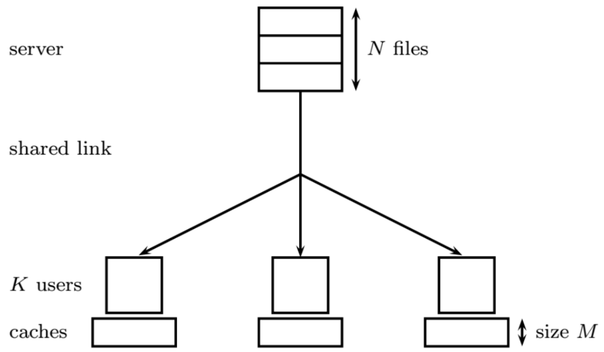

Fig. 1. Coded caching system.

Each user can access the file set $\mathcal{W}$. Let $\mathcal{Z}_k$ denote the packet subset of $\mathcal{W}$ cached by user k, where $k \in \mathcal{K}$. Note that the size of $\mathcal{Z}_k$ cannot be greater than each user's cache memory size M, i.e., $|\mathcal{Z}_k| \leq M$.

• **Delivery Phase**: Each user requests an arbitrary file in $\mathcal{W}$. The request vector is denoted by $\mathbf{d} = (d_0, d_1, \ldots, d_{K-1})$, i.e., user k requests file W_{d_k}, where $k \in \mathcal{K}$ and $d_k \in [0 : N-1]$. Once the server receives the request vector $\mathbf{d}$, it broadcasts at most RF packets such that each user can recover its requested file together with the contents in its own cache.

The above two phases can be described by the PDA that is defined below.

Definition 1 [21]: Given $K, F, Z, S \in \mathbb{N}^+$, an $F \times K$ array $\mathbf{P} = (\mathbf{P}(i,j))$, where $i \in [0 : F-1]$, $j \in [0 : K-1]$, and $\mathbf{P}(i,j) \in [0 : S-1] \cup \{*\}$, is called a (K, F, Z, S) PDA if the following conditions are satisfied:

C1. Symbol "$*$" appears exactly Z times in each column;

C2. Each integer of $[0 : S-1]$ appears at least once in the array;

C3. For any two distinct entries $\mathbf{P}(i_1, j_1)$ and $\mathbf{P}(i_2, j_2)$, $\mathbf{P}(i_1, j_1) = \mathbf{P}(i_2, j_2) = s$ is an integer only if

(a). $i_1 \neq i_2, j_1 \neq j_2$, i.e., they lie in distinct rows and distinct columns;

(b). $\mathbf{P}(i_1, j_2) = \mathbf{P}(i_2, j_1) = *$, i.e., the corresponding 2×2 subarray formed by rows i_1, i_2 and columns j_1, j_2 must be in one of the following forms

$$\begin{pmatrix} s & * \\ * & s \end{pmatrix}, \quad \begin{pmatrix} * & s \\ s & * \end{pmatrix}.$$

E.g., the following array is a (4, 4, 2, 4) PDA.

$$\mathbf{P} = \begin{pmatrix} 0 & * & 2 & * \\ * & 0 & * & 2 \\ * & 1 & * & 3 \\ 1 & * & 3 & * \end{pmatrix}. \tag{1}$$

Algorithm 1 has been introduced to realize the PDA based coded caching schemes in [21]. Given a (K, F, Z, S) PDA $\mathbf{P}$ with columns representing the user indices and rows representing the packet indices, if $\mathbf{P}(j,k) = *$, user k has cached the jth packet of all the files. Condition C1 of *Definition 1* implies that all the users have the same memory size and the memory ratio is $\frac{M}{N} = \frac{Z}{F}$. If $\mathbf{P}(j,k) = s$, where $s \in [0 : S-1]$, the jth packet of all the files is not cached by user k. The XOR of the requested packets indicated by s will be broadcast by the server at time slot s. Condition C3 of *Definition 1* guarantees that user k can obtain its required packet, since it has cached all the other packets in the multicast message except the requested one. Finally, Condition C2 of *Definition 1* implies that the number of packets transmitted by the server is exactly S and the transmission rate is $R = \frac{S}{F}$. Furthermore, the coding gain in each time slot $s \in [0 : S-1]$, denoted by g_s, equals to the occurrences of s in $\mathbf{P}$. This is because the coded packet broadcast at time slot s is useful for g_s users. Based on Algorithm 1, the following lemma can be obtained.

Algorithm 1 Coded Caching Scheme Based on PDA [21]

Procedure Placement ($\mathbf{P}$, $\mathcal{W}$)
    Split each file $W_n \in \mathcal{W}$ into F packets as $W_n = \{W_{n,j} \mid j \in [0 : F-1]\}$.
    For $k \in \mathcal{K}$ **do**
        $\mathcal{Z}_k \leftarrow \{W_{n,j} \mid \mathbf{P}(j,k) = *, \forall n \in [0 : N-1]\}$;
Procedure Delivery ($\mathbf{P}, \mathcal{W}, \mathbf{d}$)
    For $s = 0, 1, \ldots, S-1$ **do**
        Server sends $\oplus_{\mathbf{P}(j,k)=s, j \in [0:F-1], k \in [0:K-1]} W_{d_k, j}$.

Lemma 1 [21]: Given a (K, F, Z, S) PDA, there always exists an F-division (K, M, N) coded caching scheme with a memory ratio of $\frac{M}{N} = \frac{Z}{F}$ and a transmission rate of $R = \frac{S}{F}$.

Therefore, based on Algorithm 1, the PDA $\mathbf{P}$ in (1) can realize a 4-division (4, 2, 4) coded caching scheme as follows.

• **Placement Phase**: Each file W_n is divided into 4 packets, i.e., $W_n = \{W_{n,0}, W_{n,1}, W_{n,2}, W_{n,3}\}$, where $n \in [0 : 3]$. The contents cached by each user are

$$\begin{aligned} \mathcal{Z}_0 &= \{W_{n,1}, W_{n,2} \mid n \in [0:3]\}; \\ \mathcal{Z}_1 &= \{W_{n,0}, W_{n,3} \mid n \in [0:3]\}; \\ \mathcal{Z}_2 &= \{W_{n,1}, W_{n,2} \mid n \in [0:3]\}; \\ \mathcal{Z}_3 &= \{W_{n,0}, W_{n,3} \mid n \in [0:3]\}. \end{aligned}$$

• **Delivery Phase**: Assume that the request vector is $\mathbf{d} = (0, 1, 2, 3)$. The packets sent by the server at four time slots are listed as follows. Time slot 0: $W_{0,0} \oplus W_{1,1}$; Time slot 1: $W_{0,3} \oplus W_{1,2}$; Time slot 2: $W_{2,0} \oplus W_{3,1}$; Time slot 3: $W_{2,3} \oplus W_{3,2}$. Then, each user can recover its requested file. E.g., user 0 requests file $W_0 = \{W_{0,0}, W_{0,1}, W_{0,2}, W_{0,3}\}$ and has cached $W_{0,1}$ and $W_{0,2}$. At time slot 0, it can receive $W_{0,0} \oplus W_{1,1}$. With this, $W_{0,0}$ can be recovered since it has cached $W_{1,1}$. At time slot 1, it can receive $W_{0,3} \oplus W_{1,2}$. Afterwards, it can recover $W_{0,3}$ since it has cached $W_{1,2}$. The transmission rate is $R = \frac{4}{4} = 1$.

B. Orthogonal Arrays and Proper Orthogonal Arrays

Definition 2 [31]: Given any $m, q, t \in \mathbb{N}^+$ with $q \geq 2$ and $t \leq m$, let $\mathbf{F}$ denote an $l \times m$ matrix defined over $[0 : q-1]$. It is called an orthogonal array (OA) with a strength of t, denoted as $\text{OA}_\lambda(l, m, q, t)$, if for each subset $\mathcal{S} \in \binom{[0:m-1]}{t}$ with size t, every t-length ($t \leq m$) row vector appears exactly $\lambda = \frac{l}{q^t}$ times in $\mathbf{F}|_\mathcal{S}$.

Since $l = \lambda q^t$ for any $\mathrm{OA}_\lambda(l, m, q, t)$, it can be simplified into $\mathrm{OA}_\lambda(m, q, t)$, where λ is the index of the OA. Note that if $\lambda = 1$, it can be omitted. E.g., with $m = 3, q = 2$ and $t = 2$, we can consider the following matrix.

$$\mathbf{F} = (\mathbf{f}_0; \mathbf{f}_1; \mathbf{f}_2; \mathbf{f}_3) = ((110); (000); (101); (011)). \tag{2}$$

For each $\mathcal{S} \in \binom{[0:2]}{2}$, we have

$$\begin{aligned} \mathbf{F}|_{\{0,1\}} &= ((11); (00); (10); (01)); \\ \mathbf{F}|_{\{1,2\}} &= ((10); (00); (01); (11)); \\ \mathbf{F}|_{\{0,2\}} &= ((10); (00); (11); (01)). \end{aligned}$$

It can be seen that $\mathbf{F}$ of (2) satisfies *Definition 2*. It is an $\mathrm{OA}(3, 2, 2)$.

Based on the definition of OA, we also need a particular type of OA. It is called the proper OA (POA), which will enable the design of the new PDAs.

Definition 3: Given any $m, q \in \mathbb{N}^+$ with $q \geq 2$ and $m \geq 2$, an $\mathrm{OA}(m, q, m-1)$ is called a proper OA, denoted by $\mathrm{POA}(m, q, m-1)$, if the sum (mod q) of each row is a constant.

Since the POAs are crucial to our construction, we need to show the existence of POAs. In fact, it is true that the POAs always exist for any integers m and q, where $m \geq 2$ and $q \geq 2$.

Lemma 2: Let $\mathbf{F}$ denote a $q^{m-1} \times m$ matrix with the set of all row vectors given as

$$\mathcal{F} = \{(f_0, f_1, \ldots, f_{m-2}, x - \sum_{i=0}^{m-2} f_i) \mid f_0, f_1, \ldots, f_{m-2} \in [0 : q-1]\}, \tag{3}$$

where $x \in [0 : q-1]$, $m \geq 2$ and $q \geq 2$, then $\mathbf{F}$ is a $\mathrm{POA}(m, q, m-1)$.

Proof: Given any subset $\mathcal{S} \in \binom{[0:m-1]}{m-1}$, if $\mathcal{S} = [0 : m-2]$, it can be seen that every $(m-1)$-length row vector appears exactly once in $\mathbf{F}|_{\mathcal{S}}$. Let us consider $\mathcal{S} = [0 : m-1] \backslash \{j\}$, where $j \in [0 : m-2]$. In order to show that every $(m-1)$-length row vector appears once in $\mathbf{F}|_{\mathcal{S}}$ for such $\mathcal{S}$, one needs to show that it is impossible for an $(m-1)$-length row vector to appear more than once in $\mathbf{F}|_{\mathcal{S}}$, since the total number of row vectors with length $(m-1)$ is q^{m-1}. Assume that there exists an $(m-1)$-length row vector appearing more than once in $\mathbf{F}|_{\mathcal{S}}$. Without loss of generality, we can assume that $\mathbf{f}|_{\mathcal{S}} = \mathbf{f}'|_{\mathcal{S}}$ and $\mathbf{f} \neq \mathbf{f}'$, where $\mathbf{f} = (f_0, f_1, \ldots, f_{m-2}, x - \sum_{i=0}^{m-2} f_i) \in \mathcal{F}$ and $\mathbf{f}' = (f'_0, f'_1, \ldots, f'_{m-2}, x - \sum_{i=0}^{m-2} f'_i) \in \mathcal{F}$. This implies that $f_j = f'_j$, i.e., $\mathbf{f} = \mathbf{f}'$, which contradicts the hypothesis. Therefore, $\mathbf{F}$ is an $\mathrm{OA}(m, q, m-1)$. Furthermore, the sum of each row of $\mathbf{F}$ equals to x. Hence, it is also a $\mathrm{POA}(m, q, m-1)$. ■

It can be seen that the matrix of (2) is a $\mathrm{POA}(3, 2, 2)$ since the sum of each row is 0. It is worth pointing out that an $\mathrm{OA}(m, q, m-1)$ is not always a $\mathrm{POA}(m, q, m-1)$. E.g., with $m = 3, q = 3$ and $t = 2$, the following matrix $\mathbf{F}$ is an $\mathrm{OA}(3, 3, 2)$. However, it is not a $\mathrm{POA}(3, 3, 2)$.

$$\mathbf{F} = ((000); (011); (022); (101); (112); (120); (202); (210); (221)). \tag{4}$$

In this paper, we only consider the OAs with all rows being different, and thus there is no need to distinguish an OA and the set of its row vectors.

C. PDA Constructions via OAs

It has been shown that row indices of PDAs can be represented by OAs [26]. We briefly review the existing constructions of PDAs based on OAs [18], [21], [25], [26].

Given any $q, m \in \mathbb{N}^+$ with $q \geq 2$, let $\mathcal{F} = \mathrm{OA}(m, q, m)$ and $\mathcal{K} = \{\mathbf{k} = (\xi_0, c_0) \mid \xi_0 \in [0 : m], c_0 \in [0 : q-1]\}$. A $q^m \times (m+1)q$ array $\mathbf{P} = (\mathbf{P}(\mathbf{f}, \mathbf{k}))$, where $\mathbf{f} = (f_0, f_1, \ldots, f_{m-1}) \in \mathcal{F}$ and $\mathbf{k} = (\xi_0, c_0) \in \mathcal{K}$, can be constructed with the entry $\mathbf{P}(\mathbf{f}, \mathbf{k})$ defined as

- When $\xi_0 \in [0 : m-1]$,

$$\mathbf{P}(\mathbf{f}, \mathbf{k}) = \begin{cases} (f_0, \ldots, c_0, \ldots, f_{m-1}, f_{\xi_0} - c_0 - 1), \\ \quad \text{if } f_{\xi_0} \neq c_0\ ; \\ *, \quad \text{otherwise.} \end{cases} \tag{5}$$

- When $\xi_0 = m$,

$$\mathbf{P}(\mathbf{f}, \mathbf{k}) = \begin{cases} (f_0, f_1, \ldots, f_{m-1}, c_0 - \sum_{i=0}^{m-1} f_i - 1), \\ \quad \text{if } \sum_{i=0}^{m-1} f_i \neq c_0\ ; \\ *, \quad \text{otherwise.} \end{cases} \tag{6}$$

Note that mod q computations are performed with the construction above. This construction leads to the following result.

Lemma 3 [21]: Given any $q, m \in \mathbb{N}^+$ with $q \geq 2$, there always exists an $((m+1)q,\ q^m,\ q^{m-1},\ q^{m+1} - q^m)$ PDA.

The PDA construction of [18] can be seen as a generalization of (5) by considering $t \geq 1$. Given any $q, m, t \in \mathbb{N}^+$ with $t < m$ and $q \geq 2$, let $\mathcal{F} = \mathrm{OA}(m, q, m)$ and $\mathcal{K} = \{\mathbf{k} = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, c_1, \ldots, c_{t-1})) \mid \{\xi_0, \xi_1, \ldots, \xi_{t-1}\} \in \binom{[0:m-1]}{t}, \xi_0 < \xi_1 < \cdots < \xi_{t-1}, c_i \in [0 : q-1], i \in [0 : t-1]\}$. A $q^m \times \binom{m}{t} q^t$ array $\mathbf{P} = (\mathbf{P}(\mathbf{f}, \mathbf{k}))$, where $\mathbf{f} = (f_0, f_1, \ldots, f_{m-1}) \in \mathcal{F}$ and $\mathbf{k} = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, c_1, \ldots, c_{t-1})) \in \mathcal{K}$, can be constructed with the entry $\mathbf{P}(\mathbf{f}, \mathbf{k})$ defined as

$$\mathbf{P}(\mathbf{f}, \mathbf{k}) = \begin{cases} (f_0, \ldots, c_h, \ldots, f_{m-1}, f_{\xi_0} - c_0 - 1, & \text{if } f_{\xi_h} \neq c_h \\ \ldots, f_{\xi_{t-1}} - c_{t-1} - 1), & \text{for } h \in [0 \\ & : t-1]; \\ *, & \text{otherwise.} \end{cases} \tag{7}$$

Based on the above construction, the following result can be obtained.

Lemma 4 [18]: Given any $q, m, t \in \mathbb{N}^+$ with $q \geq 2$ and $t < m$, there always exists an $(\binom{m}{t} q^t,\ q^m,\ q^m - q^{m-t}(q-1)^t,\ q^m(q-1)^t)$ PDA.

The PDA construction of [18] was later generalized in [25] through changing its entry rule. Given any $q, z, m, t \in \mathbb{N}^+$ with $z < q$, $q \geq 2$ and $t < m$, let

$\mathcal{F} = \{(\mathbf{f},\mathbf{g}) = ((f_0, f_1, \ldots, f_{m-1}), (g_0, g_1, \ldots, g_{t-1})) \mid (\mathbf{f},\mathbf{g}) \in \mathrm{OA}(m,q,m) \times [0 : \lfloor \frac{q-1}{q-z} \rfloor - 1]^t\}$ and $\mathcal{K} = \{\mathbf{k} = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, c_1, \ldots, c_{t-1})) \mid \{\xi_0, \xi_1, \ldots, \xi_{t-1}\} \in \binom{[0:m-1]}{t}, \xi_0 < \xi_1 < \cdots < \xi_{t-1}, c_i \in [0 : q-1], i \in [0 : t-1]\}$. A $\lfloor \frac{q-1}{q-z} \rfloor^t q^m \times \binom{m}{t} q^t$ array $\mathbf{P} = (\mathbf{P}((\mathbf{f},\mathbf{g}),\mathbf{k}))$ can be constructed with the entry $\mathbf{P}((\mathbf{f},\mathbf{g}),\mathbf{k})$ defined as

$$\mathbf{P}((\mathbf{f},\mathbf{g}),\mathbf{k}) = \begin{cases} (f_0, \ldots, c_h - g_h(q-z), & \text{if } f_{\xi_h} \notin \{c_h, c_h - 1 \\ \ldots, f_{m-1}, f_{\xi_0} - c_0 - 1, & , \ldots, c_h - (z-1)\} \\ \ldots, f_{\xi_{t-1}} - c_{t-1} - 1), & \text{for } h \in [0 : t-1]; \\ *, & \text{otherwise.} \end{cases} \tag{8}$$

This construction leads to the following result with more flexible parameters.

Lemma 5 [25]: Given any $q, z, m, t \in \mathbb{N}^+$ with $q \geq 2, z < q$ and $t < m$, there always exists an $(\binom{m}{t} q^t, \lfloor \frac{q-1}{q-z} \rfloor^t q^m, \lfloor \frac{q-1}{q-z} \rfloor^t [q^m - q^{m-t}(q-z)^t], q^m(q-z)^t)$ PDA.

The above PDA constructions were further improved in [26] with a smaller subpacketization level. Given any $q, m, t \in \mathbb{N}^+$ with $q \geq 2$ and $t < m$, let $\mathcal{F} = \mathrm{OA}(m, q, m-1)$ and $\mathcal{K} = \{(\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, c_1, \ldots, c_{t-1})) \mid \{\xi_0, \xi_1, \ldots, \xi_{t-1}\} \in \binom{[0:m-1]}{t}, \xi_0 < \xi_1 < \cdots < \xi_{t-1}, c_i \in [0 : q-1], i \in [0 : t-1]\}$. An $|\mathcal{F}| \times |\mathcal{K}|$ array $\mathbf{P} = (\mathbf{P}(\mathbf{f},\mathbf{k}))$, where $\mathbf{f} = (f_0, f_1, \ldots, f_{m-1}) \in \mathcal{F}$ and $\mathbf{k} = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, c_1, \ldots, c_{t-1})) \in \mathcal{K}$, can be obtained with the entry $\mathbf{P}(\mathbf{f},\mathbf{k})$ defined as

$$\mathbf{P}(\mathbf{f},\mathbf{k}) = \begin{cases} (f_0, \ldots, c_h, \ldots, f_{m-1}, & \text{if } f_{\xi_h} \neq c_h \text{ for } h \in \\ o((f_0, \ldots, c_h, \ldots, f_{m-1}))), & [0 : t-1]; \\ *, & \text{otherwise.} \end{cases} \tag{9}$$

Note that $o((f_0, \ldots, c_h, \ldots, f_{m-1}))$ is the occurrence order of the vector $(f_0, \ldots, c_h, \ldots, f_{m-1})$ in column $\mathbf{k}$. The construction leads to the following improved result.

Lemma 6 [26]: Given any $q, m, t \in \mathbb{N}^+$ with $q \geq 2$ and $t < m$, there always exists an $(\binom{m}{t} q^t, q^{m-1}, q^{m-1} - q^{m-t-1}(q-1)^t, q^{m-1}(q-1)^t)$ PDA.

III. The New PDA Schemes

This section proposes some new PDA constructions via POAs. The designs can improve the subpacketization level or the transmission rate of some existing schemes. Our design intuition is illustrated as follows.

A. Design Intuition

Given a (K, F, Z, S) PDA $\mathbf{P}'$, if we replace Z_0 integer entries in each column of $\mathbf{P}'$ by "$*$"s, the resulting array is a $(K, F, Z + Z_0, S_0)$ PDA, denoted by $\mathbf{P}_0$. Note that $S_0 \leq S$ always holds. This implies that the transmission rate of the scheme based on $\mathbf{P}_0$ may be smaller than the scheme based on $\mathbf{P}'$. Furthermore, we prefer to design a PDA $\mathbf{P} = (\mathbf{P}_0; \mathbf{P}_1)$

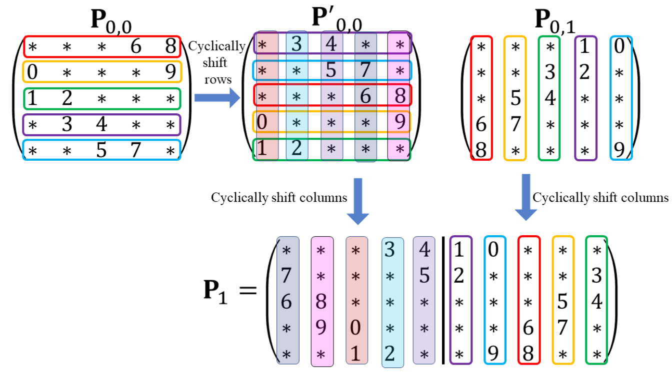

Fig. 2. Procedure of generating $\mathbf{P}_1$ from $\mathbf{P}_0$.

with the same memory ratio as $\mathbf{P}_0$ by adding a well designed $F_0 \times K$ array $\mathbf{P}_1$ without increasing S_0. This is because the new transmission rate $\frac{S_0}{F+F_0}$ is smaller than $\frac{S_0}{F}$. The following example illustrates the main idea of our construction. Given the following $(10, 5, 1, 20)$ PDA $\mathbf{P}'$,

$$\mathbf{P}' = \left(\begin{array}{ccccc|ccccc} * & 12 & 14 & 6 & 8 & * & 11 & 10 & 1 & 0 \\ 0 & * & 15 & 16 & 9 & 12 & 13 & 3 & 2 & * \\ 1 & 2 & * & 17 & 18 & 14 & 5 & 4 & * & 15 \\ 10 & 3 & 4 & * & 19 & 6 & 7 & * & 17 & 16 \\ 11 & 13 & 5 & 7 & * & 8 & * & 19 & 18 & 9 \end{array}\right),$$

if we replace its integers from 10 to 19 by "$*$"s, a new array $\mathbf{P}_0 = (\mathbf{P}_{0,0}\ \mathbf{P}_{0,1})$ can be obtained as follows.

$$\mathbf{P}_0 = \begin{array}{c} \qquad \mathbf{P}_{0,0} \qquad\qquad\qquad \mathbf{P}_{0,1} \\ \left(\begin{array}{ccccc|ccccc} * & * & * & 6 & 8 & * & * & * & 1 & 0 \\ 0 & * & * & * & 9 & * & * & 3 & 2 & * \\ 1 & 2 & * & * & * & * & 5 & 4 & * & * \\ * & 3 & 4 & * & * & 6 & 7 & * & * & * \\ * & * & 5 & 7 & * & 8 & * & * & * & 9 \end{array}\right) \end{array}.$$

Now we design an appropriate array $\mathbf{P}_1$ by adjusting the integer entry of $\mathbf{P}_0$, as shown in Fig.2. $\mathbf{P}_0$ is divided into two 5×5 arrays, i.e., $\mathbf{P}_0 = (\mathbf{P}_{0,0}\ \mathbf{P}_{0,1})$. Then, from top to bottom, each row of $\mathbf{P}_{0,0}$ is cyclically shifted by two positions, resulting in array $\mathbf{P}'_{0,0}$. From left to right, each column of $\mathbf{P}'_{0,0}$ is further cyclically shifted by two positions, resulting in the left half of $\mathbf{P}_1$. Finally, from left to right, each column of $\mathbf{P}_{0,1}$ is cyclically shifted by two positions, resulting in the right half of $\mathbf{P}_1$. Adding the well designed array $\mathbf{P}_1$ to $\mathbf{P}_0$ from top to bottom, a new $(10, 10, 6, 10)$ PDA $\mathbf{P} = (\mathbf{P}_0; \mathbf{P}_1)$ can be obtained as follows.

$$\mathbf{P} = \left(\begin{array}{ccccc|ccccc} * & * & * & 6 & 8 & * & * & * & 1 & 0 \\ 0 & * & * & * & 9 & * & * & 3 & 2 & * \\ 1 & 2 & * & * & * & * & 5 & 4 & * & * \\ * & 3 & 4 & * & * & 6 & 7 & * & * & * \\ * & * & 5 & 7 & * & 8 & * & * & * & 9 \\ \hline * & * & * & 3 & 4 & 1 & 0 & * & * & * \\ 7 & * & * & * & 5 & 2 & * & * & * & 3 \\ 6 & 8 & * & * & * & * & * & * & 5 & 4 \\ * & 9 & 0 & * & * & * & * & 6 & 7 & * \\ * & * & 1 & 2 & * & * & 9 & 8 & * & * \end{array}\right).$$

In general, a PDA $\mathbf{P} = (\mathbf{P}_0; \mathbf{P}_1; \ldots; \mathbf{P}_L)$ constructed by the above method can be viewed as replacing the same number of integer entries by "$*$"s in each column of a given baseline array $\mathbf{P}'$, and further adding new arrays $\mathbf{P}_1, \ldots, \mathbf{P}_L$ with the same memory ratio as $\mathbf{P}_0$ from top to bottom. In order to minimize the transmission rate of the scheme based on $\mathbf{P}$, one needs to guarantee if an integer entry s occurring in some row and column of $\mathbf{P}'$ is replaced by "$*$", all the entries of $\mathbf{P}'$ containing s are also replaced by "$*$"s as well. Furthermore, the newly added arrays $\mathbf{P}_1, \ldots, \mathbf{P}_L$ should be well designed such that their integer entries are the same as those in $\mathbf{P}_0$. This implies that the main technical challenge for the above construction is how to design a baseline array $\mathbf{P}'$ and the newly added arrays $\mathbf{P}_1, \ldots, \mathbf{P}_L$ that can satisfy such constraints.

It should be pointed out that our proposed method of PDA construction is different with the construction of [25]. In [25], all row indices of the newly added arrays are generated by the same OA(m, q, m) and their integer entries are obtained by moving the entries of a designed array in a counter clockwise manner. Furthermore, if the technique of [25] is directly applied to that of [26], the above constraints will be violated. Therefore, in order to obtain a PDA with the largest possible coding gain for each integer entry, some more empirical insights and technical delicacies should be utilized so that the above constraints can be satisfied.

B. The New PDA Construction

Based on the above observation, this subsection proposes a novel framework of constructing PDA. It can be considered as a generalization of the work of [26]. However, it requires a more delicate selection of the row indices of arrays. Unlike the construction of [25], i.e., the row indices of the PDA are obtained by reusing the row indices of the PDA (in [18]) $\lfloor \frac{q-1}{q-z} \rfloor^t$ times. Our construction is realized by carefully selecting the POAs. Now let us present some useful notations for our construction.

- Given any $q, z, m, t \in \mathbb{N}^+$ with $z < q$ and $t < m$, let $\mathcal{E} = \{(g_0, g_1, \ldots, g_{t-1}) \mid g_i \in [0 : \lfloor \frac{q-1}{q-z} \rfloor - 1], i \in [0 : t-1]\}$, and let

$$\begin{aligned}\mathbf{F}_{\mathbf{g}_j} &= (\mathbf{f}_0^{(j)}; \mathbf{f}_1^{(j)}; \ldots; \mathbf{f}_{q^{m-1}-1}^{(j)}) \\ &= ((f_{0,0}^{(j)}, f_{0,1}^{(j)}, \ldots, f_{0,m-1}^{(j)}); (f_{1,0}^{(j)}, f_{1,1}^{(j)}, \ldots, f_{1,m-1}^{(j)}); \ldots; \\ &\quad (f_{q^{m-1}-1,0}^{(j)}, f_{q^{m-1}-1,1}^{(j)}, \ldots, f_{q^{m-1}-1,m-1}^{(j)}))\end{aligned} \tag{10}$$

denote a POA$(m, q, m-1)$ such that $\sum_{r=0}^{m-1} f_{s,r}^{(j)} = x(q-z)$ for $s \in [0 : q^{m-1} - 1]$, where $x = \sum_{i=0}^{t-1} g_i^{(j)}$, $\mathbf{g}_j = (g_0^{(j)}, g_1^{(j)}, \ldots, g_{t-1}^{(j)}) \in \mathcal{E}$ and $\mathcal{E} = \{\mathbf{g}_0, \mathbf{g}_1, \ldots, \mathbf{g}_{\lfloor \frac{q-1}{q-z} \rfloor^t - 1}\}$.

- Let $\mathscr{I} = \{\{\xi_0, \xi_1, \ldots, \xi_{t-1}\} \mid \{\xi_0, \xi_1, \ldots, \xi_{t-1}\} \in \binom{[0:m-1]}{t}, 0 \le \xi_0 < \xi_1 < \cdots < \xi_{t-1} < m\}$ and $\mathcal{C} = \{(c_0, c_1, \ldots, c_{t-1}) \mid c_i \in [0 : q-1], i \in [0 : t-1]\}$.

Armed with the above notations, the new PDA construction can be proposed as follows.

Construction 1: Given any $q, z, m, t \in \mathbb{N}^+$ with $z < q$ and $t < m$, let $\mathcal{K} = \{(\mathcal{I}, \mathbf{c}) = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, \ldots, c_{t-1})) \mid \mathcal{I} \in \mathscr{I}, \mathbf{c} \in \mathcal{C}\}$ and $\mathcal{F} = \bigcup_{j \in [0:\lfloor \frac{q-1}{q-z} \rfloor^t - 1]} \mathcal{F}_{\mathbf{g}_j}^{(j)}$, where $\mathcal{F}_{\mathbf{g}_j}^{(j)} = \{(\mathbf{f}_s^{(j)}, \mathbf{g}_j) = ((f_{s,0}^{(j)}, \ldots, f_{s,m-1}^{(j)}), (g_0^{(j)}, \ldots, g_{t-1}^{(j)})) \mid s \in [0 : q^{m-1} - 1]\}$ and $\mathbf{f}_s^{(j)} \in \mathbf{F}_{\mathbf{g}_j}$. An $F \times K$ array $\mathbf{P} = (\mathbf{P}_0; \ldots; \mathbf{P}_j; \ldots; \mathbf{P}_{\lfloor \frac{q-1}{q-z} \rfloor^t - 1})$ can be constructed with the entries of $\mathbf{P}_j = (\mathbf{P}_j((\mathbf{f}_s^{(j)}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c})))$ defined as

$$\begin{aligned}&\mathbf{P}_j((\mathbf{f}_s^{(j)}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c})) \\ &= \begin{cases} (\mathbf{v}, o(\mathbf{v})), & \text{if } f_{s,\xi_h}^{(j)} \notin \{c_h, c_h - 1, \ldots \\ & \quad , c_h - (z-1)\} \text{ for } h \in \\ & \quad [0 : t-1]; \\ *, & \text{otherwise,} \end{cases}\end{aligned} \tag{11}$$

where $(\mathbf{f}_s^{(j)}, \mathbf{g}_j) \in \mathcal{F}_{\mathbf{g}_j}^{(j)}$, $(\mathcal{I}, \mathbf{c}) \in \mathcal{K}$, and $\mathbf{v} = (v_0, v_1, \ldots, v_{m-1}) \in [0 : q-1]^m$ such that

$$v_i = \begin{cases} c_h - g_h^{(j)}(q-z), & \text{if } i = \xi_h, h \in [0 : t-1]; \\ f_{s,i}^{(j)}, & \text{otherwise.} \end{cases} \tag{12}$$

Note that $o(\mathbf{v})$ is the occurrence order of vector $\mathbf{v}$ in column $(\mathcal{I}, \mathbf{c})$ and the computations are performed in mod q.

The following example illustrates the above construction.

Example 1: Given $m = 2$, $q = 5$ and $t = 1$, when $z = 3$, we have $\lfloor \frac{q-1}{q-z} \rfloor = 2$ and $\mathcal{E} = \{(0), (1)\}$. Let $\mathbf{F}_{(0)}$ and $\mathbf{F}_{(1)}$ denote two POA$(2, 5, 1)$s that are defined below. Notice that

$$\begin{aligned}\mathbf{F}_{(0)} &= ((00); (14); (23); (32); (41)); \\ \mathbf{F}_{(1)} &= ((02); (11); (20); (34); (43)).\end{aligned}$$

the sum of each row of $\mathbf{F}_{(0)}$ is $0 \times 2 = 0$ and the sum of each row of $\mathbf{F}_{(1)}$ is $1 \times 2 = 2$. Hence, we have

$$\begin{aligned}\mathcal{F} &= (\mathbf{F}_{(0)} \times \{(0)\}) \cup (\mathbf{F}_{(1)} \times \{(1)\}); \\ \mathcal{K} &= \{(\{\xi_0\}, (c_0)) \mid \xi_0 \in [0 : 1], c_0 \in [0 : 4]\}.\end{aligned}$$

For $(\{\xi_0\}, (c_0)) = (\{0\}, (0)) \in \mathcal{K}$, we have $\{c_0, c_0 - 1, \ldots, c_0 - (z-1)\} = \{0, 3, 4\}$. Based on (11) and (12), for any $((f_{s,0}^{(j)}, f_{s,1}^{(j)}), (g_0^{(j)})) \in \mathcal{F}$, we have $\mathbf{P}(((f_{s,0}^{(j)}, f_{s,1}^{(j)}), (g_0^{(j)})), (\{0\}, (0))) = *$, if $f_{s,0}^{(j)} \in \{0, 3, 4\}$; and $\mathbf{P}(((f_{s,0}^{(j)}, f_{s,1}^{(j)}), (g_0^{(j)})), (\{0\}, (0))) = (-2g_0^{(j)}, f_{s,1}^{(j)})$, if $f_{s,0}^{(j)} \notin \{0, 3, 4\}$. Furthermore, based on (11), if $f_{s,0}^{(j)} \notin \{0, 3, 4\}$, $\mathbf{P}(((f_{s,0}^{(j)}, f_{s,1}^{(j)}), (g_0^{(j)})), (\{0\}, (0))) = (-2g_0^{(j)}, f_{s,1}^{(j)}, o((-2g_0^{(j)}, f_{s,1}^{(j)})))$. E.g., since (04) first occurs in column $(\{0\}, (0))$, and with the assumption that the occurrence order starts from 0, it can be seen that $\mathbf{P}(((14), (0)), (\{0\}, (0))) = (040)$. Similarly, we can obtain $\mathbf{P}(((f_{s,0}^{(j)}, f_{s,1}^{(j)}), (g_0^{(j)})), (\{\xi_0\}, (c_0)))$ for any $((f_{s,0}^{(j)}, f_{s,1}^{(j)}), (g_0^{(j)})) \in \mathcal{F}$ and $(\{\xi_0\}, (c_0)) \in \mathcal{K}$. As a result, the PDA $\mathbf{P} = (\mathbf{P}_0; \mathbf{P}_1)$ can be obtained as follows.

$(10, 10, 6, 10)$ PDA $\mathbf{P}$

$(\mathbf{f}, \mathbf{g}_j)/$	$\{0\}$					$\{1\}$				
$(\mathcal{I}, \mathbf{c})$	(0)	(1)	(2)	(3)	(4)	(0)	(1)	(2)	(3)	(4)
((00),(0))	$*$	$*$	$*$	(300)	(400)	$*$	$*$	$*$	(030)	(040)
((14),(0))	(040)	$*$	$*$	$*$	(440)	$*$	$*$	(120)	(130)	$*$
((23),(0))	(030)	(130)	$*$	$*$	$*$	$*$	(210)	(220)	$*$	$*$
((32),(0))	$*$	(120)	(220)	$*$	$*$	(300)	(310)	$*$	$*$	$*$
((41),(0))	$*$	$*$	(210)	(310)	$*$	(400)	$*$	$*$	$*$	(440)
((02),(1))	$*$	$*$	$*$	(120)	(220)	(030)	(040)	$*$	$*$	$*$
((11),(1))	(310)	$*$	$*$	$*$	(210)	(130)	$*$	$*$	$*$	(120)
((20),(1))	(300)	(400)	$*$	$*$	$*$	$*$	$*$	$*$	(210)	(220)
((34),(1))	$*$	(440)	(040)	$*$	$*$	$*$	$*$	(300)	(310)	$*$
((43),(1))	$*$	$*$	(030)	(130)	$*$	$*$	(440)	(400)	$*$	$*$

Based on *Construction 1*, the following result presents a new class of PDAs that can yield a coded caching scheme with a smaller subpacketization level than the one in [25]. For the sake of presentation consistency, its proof is given in Appendix A.

Theorem 1: Given any $q, z, m, t \in \mathbb{N}^+$ with $q \geq 2, z < q$ and $t < m$, there always exists an $(\binom{m}{t}q^t, \lfloor \frac{q-1}{q-z} \rfloor^t q^{m-1}, \lfloor \frac{q-1}{q-z} \rfloor^t [q^{m-1} - q^{m-t-1}(q-z)^t], q^{m-1}(q-z)^t)$ PDA which yields a $\lfloor \frac{q-1}{q-z} \rfloor^t q^{m-1}$-division $(\binom{m}{t}q^t, M, N)$ coded caching scheme with a memory ratio of

$$\frac{M}{N} = 1 - (\frac{q-z}{q})^t$$

and a transmission rate of

$$R = \frac{(q-z)^t}{\lfloor \frac{q-1}{q-z} \rfloor^t}.$$

C. The Generic Transform

Inspired by *Construction 1*, we now introduce an effective transform of *Construction 1* that can derive a coded caching scheme with a smaller subpacketization level than the scheme in *Theorem 1*. The transform can be performed through the following two steps.

Step 1: Choose the new $\hat{\mathbf{P}}_0, \hat{\mathbf{P}}_1, \ldots, \hat{\mathbf{P}}_{\lfloor \frac{q-1}{q-z} \rfloor^t - 1}$ as the baseline arrays.

Let $\mathbf{P} = (\mathbf{P}_0; \ldots; \mathbf{P}_j; \ldots; \mathbf{P}_{\lfloor \frac{q-1}{q-z} \rfloor^t - 1})$ denote an array generated from *Construction 1* with the row indices of $\mathbf{P}_j$ indexed by $\mathcal{F}^{(j)}_{\mathbf{g}_j}$, where the first $m-1$ coordinates of the row indices of $\mathbf{P}_j$ (the first $m-1$ coordinates of the rows of $\mathbf{F}_{\mathbf{g}_j}$) are arranged in the lexicographic order from top to bottom. For any $j \in [1 : \lfloor \frac{q-1}{q-z} \rfloor^t - 1]$, array $\mathbf{P}'_j$ can be obtained from $\mathbf{P}_j$ as follows.

- Let $\mathbf{P}'_0 = \mathbf{P}_0$. Given any $\mathbf{g}_j = (g^{(j)}_0, g^{(j)}_1, \ldots, g^{(j)}_{t-1}) \in \mathcal{E}$ and $\mathbf{g}_0 = (0, 0, \ldots, 0)$, if an entry $\mathbf{e} = (\mathbf{v}, o(\mathbf{v}))$ appears in row $((f^{(0)}_{s0}, f^{(0)}_{s1}, \ldots, f^{(0)}_{s\xi_i}, \ldots, f^{(0)}_{sm-1}), \mathbf{g}_0)$ and column $(\mathcal{I}, \mathbf{c}) = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, \mathbf{c})$ of $\mathbf{P}_0$, the entry of $\mathbf{P}_j$ that contains $\mathbf{v}$ in row $((f^{(0)}_{s0}, \ldots, f^{(0)}_{s\xi_0} + g^{(j)}_0(q-z), \ldots, f^{(0)}_{s\xi_1} + g^{(j)}_1(q-z), \ldots, f^{(0)}_{s\xi_{t-1}} + g^{(j)}_{t-1}(q-z), \ldots, f^{(0)}_{sm-1}), \mathbf{g}_j)$ and column $(\mathcal{I}, \mathbf{c}')$ will be modified as $\mathbf{e}$. Applying this manner to all the entries of $\mathbf{P}_0$, we can obtain a new array $\mathbf{P}'_j$ by modifying the entries of $\mathbf{P}_j$.
- Let $\hat{\mathbf{P}}_0 = \mathbf{P}'_0$. For any $j \in [1 : \lfloor \frac{q-1}{q-z} \rfloor^t - 1]$, the baseline array $\hat{\mathbf{P}}_j$ is obtained by selecting the columns of $\mathbf{P}'_j$ indexed by $(\mathcal{I}, \mathbf{c})$, where $\mathcal{I} \in \binom{[0:m-2]}{t}$ and $\mathbf{c} \in \{(c_0, c_1, \ldots, c_{t-1}) \mid c_i \in [0 : q-1], i \in [0 : t-1]\}$.

Step 2: Transform the baseline arrays into the new PDA $\hat{\mathbf{P}}$.

This transform is performed by arranging the baseline arrays from left to right as $\hat{\mathbf{P}} = (\hat{\mathbf{P}}_0 \ \ldots \ \hat{\mathbf{P}}_j \ \ldots \ \hat{\mathbf{P}}_{\lfloor \frac{q-1}{q-z} \rfloor^t - 1})$. A new array $\hat{\mathbf{P}}$ can be obtained with the number of columns being greater than that of $\mathbf{P}$, while maintaining the same number of rows as $\mathbf{P}_0$.

By applying the transform to *Construction 1*, we have the following result, while its proof can be found in Appendix B.

Theorem 2: Given any $q, z, m, t \in \mathbb{N}^+$ with $q \geq 2, z < q$ and $t < m$, there always exists an $([\binom{m-1}{t} \lfloor \frac{q-1}{q-z} \rfloor^t + \binom{m}{t} - \binom{m-1}{t}]q^t, q^{m-1}, q^{m-1} - q^{m-t-1}(q-z)^t, q^{m-1}(q-z)^t)$ PDA which yields a q^{m-1}-division $([\binom{m-1}{t} \lfloor \frac{q-1}{q-z} \rfloor^t + \binom{m}{t} - \binom{m-1}{t}]q^t, M, N)$ coded caching scheme with a memory ratio of

$$\frac{M}{N} = 1 - (\frac{q-z}{q})^t$$

and a transmission rate of

$$R = (q-z)^t.$$

The following example illustrates the above transform and its effect.

Example 2: Let us consider the array $\mathbf{P} = (\mathbf{P}_0; \mathbf{P}_1)$ of *Example 1*. Based on the above transform, we have $\hat{\mathbf{P}}_0 = \mathbf{P}_0$. Furthermore, based on $\mathbf{P}_0$, array $\mathbf{P}'_1$ can be obtained from $\mathbf{P}_1$ by modifying its entries. E.g., given $\mathbf{g}_1 = (1) \in \mathcal{E}_1$, if the entry (040) occurs in row $((14), (0))$ and column $\{\{0\}, (0)\}$ of $\mathbf{P}_0$, the entry of $\mathbf{P}_1$ contains vector (04) in row $((34), (1))$ and column $\{\{0\}, (2)\}$ will be modified as (040) since $q - z = 2$. Similarly, the other entries of $\mathbf{P}'_1$ can be obtained in the same manner. Then, $\hat{\mathbf{P}}_1$ can be obtained by selecting the columns of $\mathbf{P}'_1$ indexed by $(\{0\}, (0)), (\{0\}, (1)), (\{0\}, (2)), (\{0\}, (3))$ and $(\{0\}, (4))$. Consequently, we have the following $(15, 5, 3, 10)$ PDA $\hat{\mathbf{P}} = (\hat{\mathbf{P}}_0 \ \hat{\mathbf{P}}_1)$ which yields a 5-division $(15, M, N)$ coded caching scheme with a memory ratio of $\frac{M}{N} = \frac{3}{5}$ and a transmission rate of $R = 2$.

$\mathbf{P}'_1$

$(\mathbf{f}, \mathbf{g}_j)/$	$\{0\}$					$\{1\}$				
$(\mathcal{I}, \mathbf{c})$	(0)	(1)	(2)	(3)	(4)	(0)	(1)	(2)	(3)	(4)
((02),(1))	$*$	$*$	$*$	(120)	(220)	(030)	(040)	$*$	$*$	$*$
((11),(1))	(310)	$*$	$*$	$*$	(210)	(130)	$*$	$*$	$*$	(120)
((20),(1))	(300)	(400)	$*$	$*$	$*$	$*$	$*$	$*$	(210)	(220)
((34),(1))	$*$	(440)	(040)	$*$	$*$	$*$	$*$	(300)	(310)	$*$
((43),(1))	$*$	$*$	(030)	(130)	$*$	$*$	(440)	(400)	$*$	$*$

$\hat{\mathbf{P}}_1$

$(\mathbf{f}, \mathbf{g}_j)/$	$\{0\}$				
$(\mathcal{I}, \mathbf{c})$	(0)	(1)	(2)	(3)	(4)
((02),(1))	$*$	$*$	$*$	(120)	(220)
((11),(1))	(310)	$*$	$*$	$*$	(210)
((20),(1))	(300)	(400)	$*$	$*$	$*$
((34),(1))	$*$	(440)	(040)	$*$	$*$
((43),(1))	$*$	$*$	(030)	(130)	$*$

$\hat{\mathbf{P}} = (\hat{\mathbf{P}}_0 \ \hat{\mathbf{P}}_1)$

$*$	$*$	$*$	(300)	(400)	$*$	$*$	$*$	(030)	(040)	$*$	$*$	$*$	(120)	(220)
(040)	$*$	$*$	$*$	(440)	$*$	$*$	(120)	(130)	$*$	(310)	$*$	$*$	$*$	(210)
(030)	(130)	$*$	$*$	$*$	$*$	(210)	(220)	$*$	$*$	(300)	(400)	$*$	$*$	$*$
$*$	(120)	(220)	$*$	$*$	(300)	(310)	$*$	$*$	$*$	$*$	(440)	(040)	$*$	$*$
$*$	$*$	(210)	(310)	$*$	(400)	$*$	$*$	$*$	(440)	$*$	$*$	(030)	(130)	$*$

D. New PDA Schemes With Coded Placement

Based on the above PDAs, two new coded caching schemes with the coded placement will be proposed. In a PDA, a "$*$" is useless, if it is not contained in any subarray shown in C3-(b) of *Definition 1*. This indicates the useless "$*$"s cannot generate multicasting opportunities in the delivery phase, i.e., these "$*$"s have no advantage in reducing the transmission rate of a coded caching scheme realized by that PDA and result in a high subpacketization level. Therefore, if each column of a (K, F, Z, S) PDA has Z' useless "$*$"s, a new coded caching scheme with a smaller memory ratio and subpacketization level can be obtained by deleting these useless "$*$"s, and using an $[F, F-Z']_{q'}$ maximum distance separable code that is defined over $\mathbb{F}_{q'}$ in the placement phase, where F and $F-Z'$ denote the length and dimension of the code, respectively. More details can be seen in [27].

Lemma 7 [27]: For any (K,F,Z,S) PDA $\mathbf{P}$, if there exist Z' useless "$*$"s in each column, we can obtain an $(F-Z')$-division (K,M,N) coded caching scheme with a memory ratio of

$$\frac{M}{N}=\frac{Z-Z'}{F-Z'}$$

and a transmission rate of

$$R=\frac{S}{F-Z'}.$$

The coding gain at each time slot is the same as that of the original scheme realized by $\mathbf{P}$ and Algorithm 1.

Remark 1: Note that the operation field size q' of *Lemma 7* must be $\mathcal{O}(F)$. Therefore, the size of each packet of files must approximate to $\log_2 F$ bits. This implies that the size of files in the server must be more than $(F-Z')\log_2 F$ so that the transmission rate of $\frac{S}{F-Z'}$ can be maintained. Given a (K,F,Z,S) PDA, $\frac{Z}{F} > \frac{Z-Z'}{F-Z'}$ and $F > F-Z'$ always hold for any $Z, F, Z' \in \mathbb{N}^+$. This also implies that we can obtain a coded caching scheme with a smaller subpacketization level and memory ratio by deleting the useless "$*$"s.

Note that there exist some useless "$*$"s in each column of the array $\mathbf{P}$ generated from *Construction 1*. Based on *Lemma 7*, the following result can be further obtained.

Theorem 3: Given any $q,m,t \in \mathbb{N}^+$ with $q \geq 2$ and $t < m$, let z_r^* denote the minimal integer in the set $\mathcal{G}_r = \{z \mid \lfloor\frac{q-1}{q-z}\rfloor = r, z \in [1:q-1]\}$, where $r \in [1:q-1]$. There exists an $(\binom{m}{t}q^t, M, N)$ coded caching scheme with a memory ratio of

$$\frac{M}{N}=\frac{1-(\frac{q-z_r^*}{q})^t}{1-(\frac{q-z_r^*}{q})^t+(\frac{q-z}{q})^t},$$

a transmission rate of

$$R=\frac{(q-z)^t}{\lfloor\frac{q-1}{q-z}\rfloor^t(1-(\frac{q-z_r^*}{q})^t+(\frac{q-z}{q})^t)},$$

and a subpacketization level of

$$F=\lfloor\frac{q-1}{q-z}\rfloor^t q^{m-1}[1+(\frac{q-z}{q})^t-(\frac{q-z_r^*}{q})^t].$$

Proof: Let $\mathbf{P}$ denote a PDA generated from *Construction 1*. When $z > z_r^*$ and $z \in \mathcal{G}_r$, array $\mathbf{P}$ can be viewed as uniformly replacing some vector entries by "$*$"s in the column where they occur from the case of $z = z_r^*$. This implies that there exist $\lfloor\frac{q-1}{q-z}\rfloor^t q^{m-1}[(\frac{q-z_r^*}{q})^t-(\frac{q-z}{q})^t]$ useless "$*$"s in each column of $\mathbf{P}$. Further based on *Theorem 1*, the number of distinct vector entries in $\mathbf{P}$ is $q^{m-1}(q-z)^t$. Combined with *Lemma 7*, the conclusion can be reached. ∎

The following example illustrates the realization of the scheme in *Theorem 3*.

Example 3: Given $m=2$, $q=5$ and $t=1$, based on *Construction 1*, when $z=1$, we obtain a PDA $\overline{\mathbf{P}}_0$ as follows.

$(10,5,1,20)$ PDA $\overline{\mathbf{P}}_0$

$(\mathbf{f},\mathbf{g}_j)/$ $(\mathcal{I},\mathbf{c})$	$\{0\}$					$\{1\}$				
	(0)	(1)	(2)	(3)	(4)	(0)	(1)	(2)	(3)	(4)
((00),(0))	$*$	(100)	(200)	(300)	(400)	$*$	(010)	(020)	(030)	(040)
((14),(0))	(040)	$*$	(240)	(340)	(440)	(100)	(110)	(120)	(130)	$*$
((23),(0))	(030)	(130)	$*$	(330)	(430)	(200)	(210)	(220)	$*$	(240)
((32),(0))	(020)	(120)	(220)	$*$	(420)	(300)	(310)	$*$	(330)	(340)
((41),(0))	(010)	(110)	(210)	(310)	$*$	(400)	$*$	(420)	(430)	(440)

If $r=1$, we have $\mathcal{G}_1 = \{1,2\}$ and $z_1^* = 1$. By taking $z=2$, the following PDA $\overline{\mathbf{P}}_1$ can be obtained. It can be viewed as uniformly replacing some vector entries by "$*$"s in the columns where they occur from $\overline{\mathbf{P}}_0$. The symbol "$\times$" implies that the neighbouring "$*$" is useless and it can be deleted.

$(10,5,2,15)$ PDA $\overline{\mathbf{P}}_1$

$(\mathbf{f},\mathbf{g}_j)/$ $(\mathcal{I},\mathbf{c})$	$\{0\}$					$\{1\}$				
	(0)	(1)	(2)	(3)	(4)	(0)	(1)	(2)	(3)	(4)
((00),(0))	$*$	$*\times$	(200)	(300)	(400)	$*$	$*\times$	(020)	(030)	(040)
((14),(0))	(040)	$*$	$*\times$	(340)	(440)	$*\times$	(110)	(120)	(130)	$*$
((23),(0))	(030)	(130)	$*$	$*\times$	(430)	(200)	(210)	(220)	$*$	$*\times$
((32),(0))	(020)	(120)	(220)	$*$	$*\times$	(300)	(310)	$*$	$*\times$	(340)
((41),(0))	$*\times$	(110)	(210)	(310)	$*$	(400)	$*$	$*\times$	(430)	(440)

Let $\overline{\mathbf{P}}_2$ denote an array obtained by deleting the useless "$*$"s in $\overline{\mathbf{P}}_1$. Based on *Lemma 7*, array $\overline{\mathbf{P}}_2$ can realize a 4-division $(10, 2.5, 10)$ coded caching scheme as follows.

• **Placement Phase**: Each file W_n is divided into 4 packets, i.e., $W_n = \{W_{n,0}, W_{n,1}, W_{n,2}, W_{n,3}\}$ for $n \in [0:9]$. Let $W_{n,4} = W_{n,0} + W_{n,1} + W_{n,2} + W_{n,3}$. Then, the contents cached by the users are

$$\mathcal{Z}_0 = \{W_{n,0} \mid n \in [0:9]\}; \quad \mathcal{Z}_1 = \{W_{n,1} \mid n \in [0:9]\};$$
$$\mathcal{Z}_2 = \{W_{n,2} \mid n \in [0:9]\}; \quad \mathcal{Z}_3 = \{W_{n,3} \mid n \in [0:9]\};$$
$$\mathcal{Z}_4 = \{W_{n,4} \mid n \in [0:9]\}; \quad \mathcal{Z}_5 = \{W_{n,0} \mid n \in [0:9]\};$$
$$\mathcal{Z}_6 = \{W_{n,4} \mid n \in [0:9]\}; \quad \mathcal{Z}_7 = \{W_{n,3} \mid n \in [0:9]\};$$
$$\mathcal{Z}_8 = \{W_{n,2} \mid n \in [0:9]\}; \quad \mathcal{Z}_9 = \{W_{n,1} \mid n \in [0:9]\}.$$

• **Delivery Phase**: Assume that the request vector is $\mathbf{d} = (0,1,2,3,4,5,6,7,8,9)$. The packets sent by the server at fifteen time slots are listed in Table III. Then, each user can decode its requested file since each file W_n can be recovered by any 4 packets out of $\{W_{n,j} \mid j \in [0:4]\}$. E.g., user 1 first decodes the required packets $W_{1,2}, W_{1,3}$ and $W_{1,4}$ from the coded packets $W_{1,2} \oplus W_{8,1}, W_{1,3} \oplus W_{7,1}$ and $W_{1,4} \oplus W_{6,1}$, respectively. Then, $W_{1,0}$ can be obtained from $W_{1,4} = W_{1,0} + W_{1,1} + W_{1,2} + W_{1,3}$ since $W_{1,1}$ has been cached by user 1. Hence, array $\overline{\mathbf{P}}_2$ can realize a coded caching scheme with a memory ratio of $\frac{M}{N} = \frac{1}{4}$, a transmission rate of $R = \frac{15}{4}$ and a subpacketization level of $F = 4$, which is in line with *Theorem 3*.

Based on *Lemma 7*, the scheme in *Theorem 2* can also be further improved with a smaller memory ratio and subpacketization level. As the proof of the following result is similar to that of *Theorem 3*, it is omitted.

Theorem 4: Given any $q,m,t \in \mathbb{N}^+$ with $q \geq 2$ and $t < m$, let z_r^* denote the minimal integer in the set $\mathcal{G}_r = \{z \mid \lfloor\frac{q-1}{q-z}\rfloor = r, z \in [1:q-1]\}$, where $r \in [1:q-1]$. There exists an $([\binom{m-1}{t}\lfloor\frac{q-1}{q-z}\rfloor^t + \binom{m}{t} - \binom{m-1}{t}]q^t, M, N)$ coded caching

TABLE III
DELIVERY STEPS IN EXAMPLE 3

Time Slot	Transmitted Signal
(040)	$W_{0,1} \oplus W_{9,0}$
(030)	$W_{0,2} \oplus W_{8,0}$
(020)	$W_{0,3} \oplus W_{7,0}$
(130)	$W_{1,2} \oplus W_{8,1}$
(120)	$W_{1,3} \oplus W_{7,1}$
(110)	$W_{1,4} \oplus W_{6,1}$
(200)	$W_{2,0} \oplus W_{5,2}$
(220)	$W_{2,3} \oplus W_{7,2}$
(210)	$W_{2,4} \oplus W_{6,2}$
(300)	$W_{3,0} \oplus W_{5,3}$
(340)	$W_{3,1} \oplus W_{9,3}$
(310)	$W_{3,4} \oplus W_{6,3}$
(400)	$W_{4,0} \oplus W_{5,4}$
(440)	$W_{4,1} \oplus W_{9,4}$
(430)	$W_{4,2} \oplus W_{8,4}$

scheme with a memory ratio of

$$\frac{M}{N} = \frac{1-(\frac{q-z_r^*}{q})^t}{1-(\frac{q-z_r^*}{q})^t+(\frac{q-z}{q})^t},$$

a transmission rate of

$$R = \frac{(q-z)^t}{1-(\frac{q-z_r^*}{q})^t+(\frac{q-z}{q})^t},$$

and a subpacketization level of

$$F = q^{m-1}[1+(\frac{q-z}{q})^t-(\frac{q-z_r^*}{q})^t].$$

IV. ANALYSES OF THE NEW PDA SCHEMES

This section further analyzes the proposed PDA schemes, verifying the results of Table II and meanwhile comparing with the existing schemes summarized in Table I.

A. Comparison of the Schemes in Theorems 1, 3, and [1], [14], [22], [24], [25], [26]

We first consider the scheme in *Theorem 1*. Similar to the discussion of [25], let q, t and z be three fixed integers. If m approaches infinity, R and $\frac{M}{N}$ will be independent with the number of users K. Since $(\frac{m}{t})^t < \binom{m}{t} < (\frac{em}{t})^t$, we have $\frac{tK^{\frac{1}{t}}}{eq} < m < \frac{tK^{\frac{1}{t}}}{q}$, where e is the Euler's number and $K = \binom{m}{t}q^t$. This implies that $F = \lfloor\frac{q-1}{q-z}\rfloor^t q^{m-1} = \mathcal{O}(\lfloor\frac{q-1}{q-z}\rfloor^t q^{\frac{tK^{\frac{1}{t}}}{q}-1})$. Therefore, when $t \geq 2$, the subpacketization level F grows sub-exponentially with the number of users K. Let $\frac{M}{N} = 1-(\frac{q-z}{q})^t$ and $K = \binom{m}{t}q^t$, the MN scheme can be obtained with a transmission rate of $R_{\mathrm{MN}} = \frac{1}{(\frac{q}{q-z})^t+\frac{1}{K}(\frac{q}{q-z})^t-1} \geq (\frac{q-z}{q})^t$. Hence, the ratio of the transmission rates between the scheme in *Theorem 1* and the MN scheme is $\frac{R}{R_{\mathrm{MN}}} \leq \frac{q^t}{\lfloor\frac{q-1}{q-z}\rfloor^t}$. This indicates that the transmission rate of the scheme in *Theorem 1* is at most $\frac{q^t}{\lfloor\frac{q-1}{q-z}\rfloor^t}$ times of that of the MN scheme.

Note that the expressions of the memory ratio, transmission rate, and subpacketization level are complex for the schemes in *Theorems 2-4*. It is challenging to yield a conclusive analysis

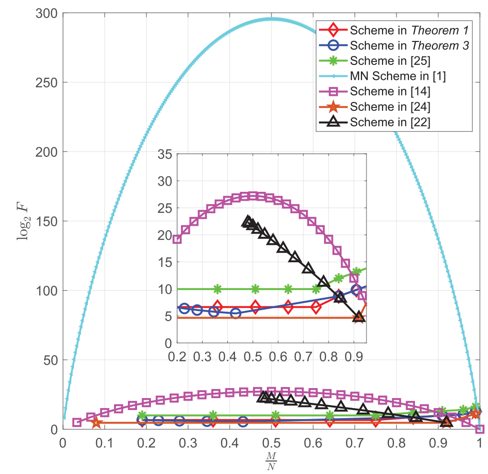

Fig. 3. Subpacketization level comparison between the schemes in *Theorems 1, 3* and [1], [14], [22], [24], [25], where $K = 300$.

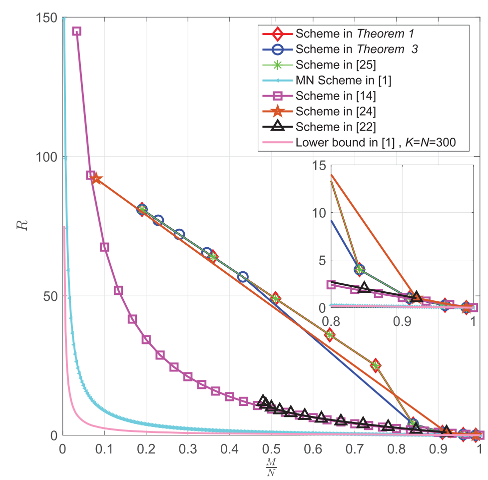

Fig. 4. Transmission rate comparison between the schemes in *Theorems 1, 3* and [1], [14], [22], [24], [25], where $K = 300$.

as above. Alternatively, we numerically compare our proposed schemes with the existing ones. When $z = 1$, our proposed scheme in *Theorem 1* will be the same as that of [26]. However, it generalizes the scheme of [26], yielding a more flexible memory size. From the parameters of the schemes in [14], [16], [17], [22], [25], and [24] and *Theorems 1-4*, it is difficult to compare them with the same number of users K. Instead, we could carefully unify the number of users K to compare the schemes in *Theorems 1, 3*, and those in [1], [14], [24], [25], and [22]. Let $m = 3, t = 2, q = 10$ and $z \in [1:9]$ for the schemes in *Theorems 1,3*, and [25]; $k = 300$ and $t \in [1:299]$ for the MN scheme in [1]; $c = 10$, $k = 30$ and $t \in [1:29]$ for the scheme in [14]; $k = 25$ and $t \in \{1,2,22,23\}$ for the scheme in [24]; $m = 25, a = 2, \lambda = 1$ and $b \in [1:12]$ for the scheme in [22]. Their subpacketization level F, memory ratio $\frac{M}{N}$, and transmission rate R can be characterized as shown in Figs. 3 and 4.

It can be seen that both the transmission rate and subpacketization level of the schemes in *Theorems 1* and *3*

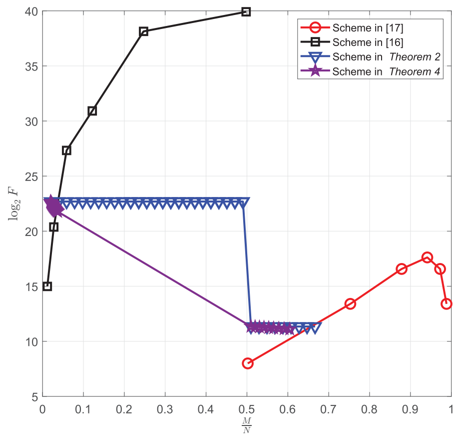

Fig. 5. Subpacketization level comparison between the schemes in *Theorems 2, 4* and [16], [17], where $K = 255$.

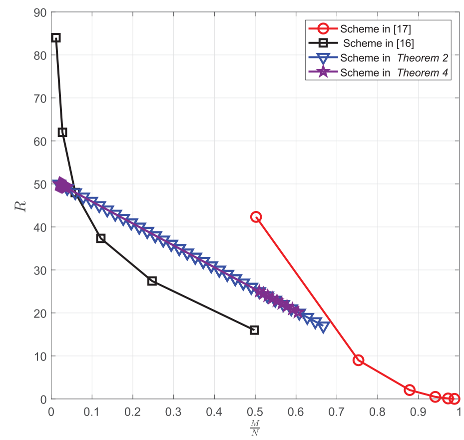

Fig. 6. Transmission rate comparison between the schemes in *Theorems 2, 4* and [16], [17], where $K = 255$.

are closed to those of the scheme in [24], which is able to achieve the minimum subpacketization level for a fixed transmission rate. This implies that the schemes characterized by *Theorems 1* and *3* also yield a good performance. Since our proposed schemes can yield a more flexible memory size, they have a wider range of applications than that of [24]. It can also be seen that with the same number of users, memory ratio and transmission rate, the scheme in *Theorem 1* has a smaller subpacketization level than that of [25]. The scheme in *Theorem 3* can even achieve a smaller subpacketization level, while maintaining almost the same transmission rate as the schemes in [25] and *Theorem 1*. When comparing with the schemes of [1], [14], and [22], our proposed schemes in *Theorems 1* and *3* have an advantage in the subpacketization level, but they are at the cost of some transmission rate.

B. Comparison of the Schemes in Theorems 2, 4 and [16], [17], [18]

We further compare the performance of our proposed schemes in *Theorems 2,4* and those in [17] and [16]. For illustration, let $q = 51, m = 3, t = 1$, $z \in [26 : 34]$ and $q = 51, m = 5, t = 1$, $z \in [1 : 25]$ for the schemes in *Theorems 2* and *4*; $k = 8, t = 1, q = 2$ and $m \in [1 : 6]$ for the schemes in [16] and [17]. Figs.5 and 6 compare their subpacketization level F and transmission rate R against the memory ratio $\frac{M}{N}$. It can be seen that when $0.0588 \leq \frac{M}{N} \leq 0.5000$, our proposed scheme in *Throrems 2* has a smaller subpacketization level than the scheme of [16], but it is at the cost of some transmission rate. Meanwhile, when $0.5000 \leq \frac{M}{N} \leq 0.6667$, our proposed schemes in *Throrems 2* and *4* have a smaller transmission rate than the scheme of [17], but with a slightly higher subpacketization level. Again, with a more flexible memory size, our proposed scheme in *Theorem 2* is expected to have a wider range of applications than the scheme of [16].

We further demonstrate the advantage of our proposed schemes in Tables IV and V. Note that the schemes in *Theorems 2, 4*, [16], and [17] are parameterized by (q, z, m, t), (q, z, m, t, z_r^*), (q, m, t, k) and (q, m, t, k), respectively. Table IV shows that in comparison with the scheme of [17], the proposed scheme in *Theorem 2* has a smaller subpacketization level and transmission rate. It also yields a slightly smaller memory ratio. Meanwhile, it is capable to serve more users simultaneously. Table V shows that in comparison with the scheme of [16], the proposed scheme in *Theorem 4* yields a smaller subpacketization level, a slightly smaller memory ratio, and supports a larger number of users. But they are realized at the cost of some transmission rate.

Finally, we consider the comparison between the scheme in *Theorem 2* and the one in [18]. For the scheme in *Theorem 2*, Let $z = q - 1$. Table VI shows that the scheme in *Theorem 2*

TABLE IV
COMPARISON BETWEEN THE SCHEME IN THEOREM 2 AND THE SCHEME IN [17]

Schemes	Parameters	K	$\frac{M}{N}$	R	F
(q, z, m, t) in *Theorem 2*	$(7, 5, 4, 1)$	70	0.71	2.0	343
(q, m, t, k) in [17]	$(2, 3, 1, 6)$	63	0.76	2.1	651
(q, z, m, t) in *Theorem 2*	$(13, 9, 4, 1)$	130	0.69	4.0	2197
(q, m, t, k) in [17]	$(2, 4, 1, 7)$	127	0.76	4.4	2667
(q, z, m, t) in *Theorem 2*	$(21, 11, 2, 1)$	63	0.52	10.0	21
(q, m, t, k) in [17]	$(2, 4, 1, 6)$	63	0.51	10.3	63
(q, z, m, t) in *Theorem 2*	$(17, 13, 4, 2)$	13056	0.94	16.0	4096
(q, m, t, k) in [17]	$(2, 4, 2, 8)$	10795	0.94	18.6	10795

TABLE V
COMPARISON BETWEEN THE SCHEME IN THEOREM 4 AND THE SCHEME IN [16]

Schemes	Parameters	K	$\frac{M}{N}$	R	F
(q, z, m, t, z_r^*) in *Theorem 4*	$(5, 2, 4, 2, 1)$	150	0.50	12.5	90
(q, m, t, k) in [16]	$(2, 5, 2, 8)$	127	0.50	9.1	3555770000
(q, z, m, t, z_r^*) in *Theorem 4*	$(8, 4, 4, 1, 1)$	32	0.20	6.4	680
(q, m, t, k) in [16]	$(2, 3, 2, 6)$	31	0.48	3.2	26040
(q, z, m, t, z_r^*) in *Theorem 4*	$(8, 2, 4, 2, 1)$	384	0.29	45.2	408
(q, m, t, k) in [16]	$(3, 4, 2, 7)$	364	0.33	40.5	5008860000
(q, z, m, t, z_r^*) in *Theorem 4*	$(13, 6, 5, 2, 1)$	1690	0.33	111.9	12506
(q, m, t, k) in [16]	$(3, 5, 2, 8)$	1093	0.33	104.1	1995520000000000

TABLE VI
COMPARISON BETWEEN THE SCHEME IN THEOREM 2 AND THE SCHEME IN [18]

Schemes and Parameters	K	$\frac{M}{N}$	R	F
New scheme in *Theorem 2*, any $q, z, m, t \in \mathbb{N}^+$ with $q \geq 2$, $z = q-1$ and $t < m$	$[\binom{m-1}{t}(q-1)^t + \binom{m}{t} - \binom{m-1}{t}]q^t$	$1 - \frac{1}{q^t}$	1	q^{m-1}
Scheme in [18], any m, q, $t \in \mathbb{N}^+$ with $t < m$ and $q \geq 2$	$\binom{m}{t}q^t$	$1 - \frac{1}{q^t}$	$\frac{1}{(q-1)^t}$	$(q-1)^t q^m$

can serve more users with a lower subpacketization level, but it is at the cost of some transmission rate.

V. CONCLUSION

This paper has proposed a novel construction of PDA via POAs. Based on the construction, some coded caching schemes have been obtained with a lower subpacketization level and a more flexible memory size. The first PDA scheme achieves an improved subpacketization level over the existing one of [25] with the same number of users, memory ratio and transmission rate. The second PDA scheme enables a larger number of users and a smaller subpacketization level than the first scheme when the memory ratio is greater than $\frac{1}{2}$. Moreover, based on the proposed PDAs, two new coded caching schemes with the coded placement have also been derived. Our analytical and numerical results have shown that the proposed schemes are able to achieve better subpacketization or transmission rate performances than the known coded caching schemes.

APPENDIX A
PROOF OF THEOREM 1

In order to prove *Theorem 1*, we need the following results.

Proposition 1: Array $\mathbf{P}$ generated from *Construction 1* satisfies Condition C3 of *Definition 1*.

Proof: Suppose that $\mathbf{P}((\mathbf{f}_s^{(j)}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c})) = \mathbf{P}((\mathbf{f}_{s'}^{(j')}, \mathbf{g}_{j'}), (\mathcal{I}', \mathbf{c}')) = (\mathbf{v}, o(\mathbf{v}))$, where

$$\mathbf{f}_s^{(j)} = (f_{s,0}^{(j)}, f_{s,1}^{(j)}, \ldots, f_{s,m-1}^{(j)});$$
$$\mathbf{f}_{s'}^{(j')} = (f_{s',0}^{(j')}, f_{s',1}^{(j')}, \ldots, f_{s',m-1}^{(j')});$$
$$\mathcal{I} = (\xi_0, \xi_1, \ldots, \xi_{t-1}); \mathcal{I}' = (\xi'_0, \xi'_1, \ldots, \xi'_{t-1});$$
$$\mathbf{g}_j = (g_0^{(j)}, g_1^{(j)}, \ldots, g_{t-1}^{(j)}); \mathbf{g}_{j'} = (g_0^{(j')}, g_1^{(j')}, \cdots, g_{t-1}^{(j')});$$
$$\mathbf{c} = (c_0, c_1, \ldots, c_{t-1}); \mathbf{c}' = (c'_0, c'_1, \ldots, c'_{t-1});$$
$$\mathbf{v} = (v_0, v_1, \ldots, v_{m-1}).$$

In order to prove C3-(a), i.e., each vector entry $(\mathbf{v}, o(\mathbf{v}))$ occurs in different rows and different columns, it is sufficient to prove that it is impossible for a vector $\mathbf{v}$ to appear more than once in some row. Assume that there exists a vector $\mathbf{v}$ to appear more than once in some row. Let us consider the following two cases under the condition of $(\mathbf{f}_s^{(j)}, \mathbf{g}_j) = (\mathbf{f}_{s'}^{(j')}, \mathbf{g}_{j'})$, where $s = s'$ and $j = j'$.

Case 1: $\mathcal{I} = \mathcal{I}'$. If $\mathbf{c} \neq \mathbf{c}'$, without loss of generality, let us assume that $c_0 \neq c'_0$. It follows from *Construction 1* that $v_{\xi_0} = c_0 - g_0^{(j)}(q-z) = c'_0 - g_0^{(j)}(q-z)$, i.e., $c_0 = c'_0$, which contradicts our hypothesis. This implies that $\mathbf{c} = \mathbf{c}'$, i.e., $(\mathcal{I}, \mathbf{c}) = (\mathcal{I}', \mathbf{c}')$.

Case 2: $\mathcal{I} \neq \mathcal{I}'$. There must exist two distinct integers, say α, $\alpha' \in [0 : m-1]$, satisfying $\alpha \in \mathcal{I}, \alpha \notin \mathcal{I}'$ and $\alpha' \in \mathcal{I}', \alpha' \notin \mathcal{I}$. Without loss of generality, let us assume that $\alpha = \xi_0, \alpha' = \xi'_0$. Based on *Construction 1*, we have $v_{\xi_0} = c_0 - g_0^{(j)}(q-z) = f_{s,\xi_0}^{(j)}$ and $v_{\xi'_0} = c'_0 - g_0^{(j)}(q-z) = f_{s,\xi'_0}^{(j)}$. This implies that $\mathbf{P}((\mathbf{f}_s^{(j)}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c})) = *$ since $f_{s,\xi_0}^{(j)} = c_0 - g_0^{(j)}(q-z) \in \{c_0, c_0-1, \ldots, c_0-(z-1)\}$, which contradicts $\mathbf{P}((\mathbf{f}_s^{(j)}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c})) = (\mathbf{v}, o(\mathbf{v}))$.

Based on the above argument, it can be seen that C3-(a) of *Definition 1* holds. Next we show that C3-(b) of *Definition 1* also holds.

Suppose that a vector $\mathbf{v}$ occurs in distinct rows and columns, i.e., rows $(\mathbf{f}_s^{(j)}, \mathbf{g}_j)$, $(\mathbf{f}_{s'}^{(j')}, \mathbf{g}_{j'})$ and columns $(\mathcal{I}, \mathbf{c}), (\mathcal{I}', \mathbf{c}')$. If $\mathcal{I} \neq \mathcal{I}'$, with the similar argument of Case 2, it can be seen that C3-(b) of *Definition 1* holds. It remains to show that C3-(b) of *Definition 1* holds for $\mathcal{I} = \mathcal{I}'$. To do this, one just needs to consider $\mathbf{g}_j \neq \mathbf{g}_{j'}$ since it is impossible for a vector $\mathbf{v}$ to appear in both entries $\mathbf{P}((\mathbf{f}_s^{(j)}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c}))$ and $\mathbf{P}((\mathbf{f}_{s'}^{(j')}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c}'))$. Let us assume that $\mathbf{P}((\mathbf{f}_s^{(j)}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c}')) \neq *$, then we have $f_{s,\xi_i}^{(j)} \in \{c'_i + 1, c'_i + 2, \ldots, c'_i + (q-z)\}$ for any $i \in [0 : t-1]$, i.e., there exists an integer $\gamma \in [1 : q-z]$ satisfying $f_{s,\xi_i}^{(j)} = c'_i + \gamma$. Based on *Construction 1*, we have $c'_i = c_i + (q-z)(g_i^{(j')} - g_i^{(j)})$. Hence, we obtain the following equation

$$f_{s,\xi_i}^{(j)} - c_i = (q-z)(g_i^{(j')} - g_i^{(j)}) + \gamma. \tag{13}$$

This is impossible for $g_i^{(j)} < g_i^{(j')}$ due to $q - z + 1 < (q-z)(g_i^{(j')} - g_i^{(j)}) + \gamma < q$ and $0 < f_{s,\xi_i}^{(j)} - c_i < q - z + 1$ by the fact $g_i^{(j)}, g_i^{(j')} \in [0 : \lfloor \frac{q-1}{q-z} \rfloor - 1]$ and $q < 2z$. If $g_i^{(j)} > g_i^{(j')}$, the equation of (13) can be written as $f_{s,\xi_i}^{(j)} - c_i + (q-z)(g_i^{(j)} - g_i^{(j')}) = \gamma$, which is also impossible since $q - z < f_{s,\xi_i}^{(j)} - c_i + (q-z)(g_i^{(j)} - g_i^{(j')}) < q$ and $1 \leq \gamma \leq q - z$. Therefore, we have $\mathbf{P}((\mathbf{f}_s^{(j)}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c}')) = *$. Similarly, we can also show that $\mathbf{P}((\mathbf{f}_s^{(j')}, \mathbf{g}_{j'}), (\mathcal{I}, \mathbf{c})) = *$. So C3-(b) of *Definition 1* holds. ■

Proposition 2: Given any $q, m, z, t \in \mathbb{N}^+$ with $q \geq 2, z < q$ and $t < m$, let $\mathbf{P} = (\mathbf{P}_0; \ldots; \mathbf{P}_j; \ldots; \mathbf{P}_{\lfloor \frac{q-1}{q-z} \rfloor^t - 1})$ denote a $q^{m-1} \lfloor \frac{q-1}{q-z} \rfloor^t \times \binom{m}{t} q^t$ array generated from *Construction 1* with row indices set $\mathcal{F} = \bigcup_{j \in [0 : \lfloor \frac{q-1}{q-z} \rfloor^t - 1]} \{(\mathbf{f}_s^{(j)}, \mathbf{g}_j) \mid s \in [0 : q^{m-1} - 1]\}$ and $\mathbf{f}_s^{(j)} \in \mathbf{F}_{\mathbf{g}_j}$. Then, all the columns of $\mathbf{P}_j$ have the same number of "$*$"s, i.e., $\mathbf{P}$ is a PDA. So the memory ratio of the $(\binom{m}{t} q^t, M, N)$ caching system realized by $\mathbf{P}$ is $\frac{M}{N} = 1 - (\frac{q-z}{q})^t$.

Proof: Note that an OA$(m, q, m-1)$ is also an OA$_{q^{m-t-1}}(m, q, t)$. This implies that $\mathbf{F}_{\mathbf{g}_j}$ defined in *Construction 1* is also an OA$_{q^{m-t-1}}(m, q, t)$. Given any column $(\mathcal{I}, \mathbf{c}) = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, c_1, \ldots, c_{t-1}))$ of $\mathbf{P}_j$, the number of row vectors $\mathbf{f}_s^{(j)} = (f_{s,0}^{(j)}, f_{s,1}^{(j)}, \ldots, f_{s,m-1}^{(j)})$ of

$\mathbf{F}_{\mathbf{g}_j}$ such that $f^{(j)}_{s,\xi_i} \notin \{c_i, c_i-1, \ldots, c_i-(z-1)\}$ for $i \in [0:t-1]$ is $(q-z)^t q^{m-t-1}$ since the occurences of each t-length row vector in $\mathbf{F}_{\mathbf{g}_j}$ is q^{m-t-1}. Then, by the entry rule of *Construction 1*, the number of non-star entries in each column of $\mathbf{P}_j$ is $(q-z)^t q^{m-t-1}$, i.e., the number of "$*$"s in each column of $\mathbf{P}_j$ is $q^{m-1}-(q-z)^t q^{m-t-1}$. Therefore, the memory ratio of the $(\binom{m}{t}q^t, M, N)$ caching system realized by $\mathbf{P}$ is $\frac{M}{N} = 1-(\frac{q-z}{q})^t$. ∎

Proposition 3: Given any $q, m, z, t \in \mathbb{N}^+$ with $q \geq 2, z < q$ and $t < m$, let $\mathbf{P} = (\mathbf{P}_0; \ldots; \mathbf{P}_j; \ldots; \mathbf{P}_{\lfloor \frac{q-1}{q-z}\rfloor^t - 1})$ denote a $q^{m-1}\lfloor \frac{q-1}{q-z}\rfloor^t \times \binom{m}{t} q^t$ array generated from *Construction 1* with row indices set $\mathcal{F} = \bigcup_{j\in[0:\lfloor \frac{q-1}{q-z}\rfloor^t-1]}\{(\mathbf{f}^{(j)}_s, \mathbf{g}_j) \mid s \in [0:q^{m-1}-1]\}$ and $\mathbf{f}^{(j)}_s \in \mathbf{F}_{\mathbf{g}_j}$. Then, each $\mathbf{P}_j$ in $\mathbf{P}$ satisfies

(a). Each vector $\mathbf{v}$ in $\mathbf{P}_j$ occurs in exactly $\binom{m}{t}$ columns, and each vector $\mathbf{v}$ in $\mathbf{P}_j$ occurs the same number of times in each column where $\mathbf{v}$ occurs;

(b). Each vector $\mathbf{v}$ occurring in $\mathbf{P}_i$ must occur in $\mathbf{P}_j$ for any $i, j \in [0:\lfloor \frac{q-1}{q-z}\rfloor^t - 1]$.

In order to prove *Proposition 3*, the following result is useful.

Proposition 4 [26]: Given any $m, q, z, t \in \mathbb{N}^+$ with $t \leq m$ and $z = 1$, let $\mathbf{P}_0$ denote a $q^{m-1} \times \binom{m}{t} q^t$ array obtained from (9) such that $\mathcal{F} =$OA$(m, q, m-1)$. Then, each vector $\mathbf{v}$ in $\mathbf{P}_0$ occurs in exactly $\binom{m}{t}$ columns, and each vector $\mathbf{v}$ in $\mathbf{P}_0$ occurs the same number of times in each column where $\mathbf{v}$ occurs.

Now let us prove *Proposition 3*.

Proof: (a). We first consider $\mathbf{P}_0$ with row indices generated by $\mathcal{F}^{(0)}_{\mathbf{g}_0} = \{(\mathbf{f}^{(0)}_s, \mathbf{g}_0) \mid s \in [0:q^{m-1}-1]\}$, where $\mathbf{f}^{(0)}_s \in \mathbf{F}_{\mathbf{g}_0}$, $\mathbf{g}_0 = (0, 0, \ldots, 0)$, and $\sum_{r=0}^{m-1} f^{(0)}_{s,r} = 0$ for $s \in [0:q^{m-1}-1]$. Based on *Proposition 4*, it can be seen that $\mathbf{P}_0$ satisfies the statement of (a) when $z = 1$. If $z > 1$, $\mathbf{P}_0$ can be viewed as replacing some vector entries by "$*$"s from the case of $z = 1$. Therefore, in order to prove that $\mathbf{P}_0$ satisfies the statement of (a) for any $z \in [2:q-1]$, one just needs to prove such vector entries that contain $\mathbf{v}$ are uniformly replaced by "$*$"s in the column where they occur. Given any $\mathbf{v} = (v_0, v_1, \ldots, v_{m-1})$ in $\mathbf{P}_0$, we denote the collection of rows in which $\mathbf{v}$ occurs corresponding to column $(\mathcal{I}, \mathbf{c}) = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, c_1, \ldots, c_{t-1}))$ as $\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}} = \{(\mathbf{f}^{(0)}_s, \mathbf{g}_0) = ((f^{(0)}_{s,0}, f^{(0)}_{s,1}, \ldots, f^{(0)}_{s,m-1}), \mathbf{g}_0) \mid \mathbf{P}((\mathbf{f}^{(0)}_s, \mathbf{g}_0), (\mathcal{I}, \mathbf{c})) = (\mathbf{v}, o(\mathbf{v})), s \in [0:q^{m-1}-1]\}$.
Similarly, let $\mathcal{F}^{(0)}_{\mathcal{I}',\mathbf{v}} = \{(\mathbf{f}^{(0)}_{s'}, \mathbf{g}_0) = ((f^{(0)}_{s',0}, f^{(0)}_{s',1}, \ldots, f^{(0)}_{s',m-1}), \mathbf{g}_0) \mid \mathbf{P}((\mathbf{f}^{(0)}_{s'}, \mathbf{g}_0), (\mathcal{I}', \mathbf{c}')) = (\mathbf{v}, o(\mathbf{v})), s' \in [0:q^{m-1}-1]\}$ denote the collection of rows in which $\mathbf{v}$ occurs corresponding to column $(\mathcal{I}', \mathbf{c}') = (\{\xi'_0, \xi'_1 \ldots, \xi'_t\}, (c'_0, c'_1, \ldots, c'_{t-1}))$, where $|\mathcal{I} \cap \mathcal{I}'| = t-1$. Note that when $z = 1$, there exists a one to one mapping from $\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}}$ to $\mathcal{F}^{(0)}_{\mathcal{I}',\mathbf{v}}$:

$$\psi((\mathbf{f}^{(0)}_s, \mathbf{g}_0)) = (\mathbf{f}^{(0)}_{s'}, \mathbf{g}_0), (\mathbf{f}^{(0)}_s, \mathbf{g}_0) \in \mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}},\ (\mathbf{f}^{(0)}_{s'}, \mathbf{g}_0) \in \mathcal{F}^{(0)}_{\mathcal{I}',\mathbf{v}},$$

where $f^{(0)}_{s,i} = f^{(0)}_{s',i}$ for $i \in (\mathcal{I} \cap \mathcal{I}') \cup ([0:m-1]\backslash(\mathcal{I} \cup \mathcal{I}'))$. Assume that the vector $\mathbf{v}$ occurs in row $(\mathbf{f}^{(0)}_s, \mathbf{g}_0) = ((f^{(0)}_{s,0}, f^{(0)}_{s,1}, \ldots, f^{(0)}_{s,m-1}), (0, 0, \ldots, 0))$ and column $(\mathcal{I}, \mathbf{c})$, then the vector $\mathbf{v}$ must occur in column $(\mathcal{I}', \mathbf{c}')$ and a unique row $(\mathbf{f}^{(0)}_{s'}, \mathbf{g}_0) = ((f^{(0)}_{s',0}, f^{(0)}_{s',1}, \ldots, f^{(0)}_{s',m-1}), (0, 0, \ldots, 0))$ since ψ is a one to one mapping. Furthermore, assume that there exist two different integers ξ_w and $\xi'_{w'}$ such that $\mathcal{I}\backslash\mathcal{I}' = \{\xi_w\}$ and $\mathcal{I}'\backslash\mathcal{I} = \{\xi'_{w'}\}$. Based on *Construction 1*, there exist two integers β, β' such that $f^{(0)}_{s,\xi_w} = c_w + \beta$, $f^{(0)}_{s,\xi'_{w'}} = c'_{w'}$, $f^{(0)}_{s',\xi_w} = c_w$, and $f^{(0)}_{s',\xi'_{w'}} = c'_{w'} + \beta'$. Then, we obtain $c_w + \beta + c'_{w'} = f^{(0)}_{s,\xi_w} + f^{(0)}_{s,\xi'_{w'}} = f^{(0)}_{s',\xi_w} + f^{(0)}_{s',\xi'_{w'}} = c'_{w'} + \beta' + c_w$ due to $\sum_{r=0}^{m-1} f^{(0)}_{s,r} = 0$ for $s \in [0:q^{m-1}-1]$, i.e., $\beta = \beta'$ holds.

Now let us consider $z > 1$ for $\mathbf{P}_0$. Without loss of generality, suppose that a vector entry that contains $\mathbf{v}$ is replaced by "$*$" in row $(\mathbf{f}^{(0)}_s, \mathbf{g}_0)$ and column $(\mathcal{I}, \mathbf{c})$, then there must exist a vector entry that contains $\mathbf{v}$ converting into "$*$" in row $(\mathbf{f}^{(0)}_{s'}, \mathbf{g}_0)$ and column $(\mathcal{I}', \mathbf{c}')$ since $\beta' = \beta \notin [1:q-z]$. This implies that vector entries that contain $\mathbf{v}$ are uniformly replaced by "$*$"s in the column where they occur, i.e., array $\mathbf{P}_0$ satisfies the statement of (a). In the following we show that there always exists array $\mathbf{P}_j$ satisfying the statement of (a) for $j \in [1:\lfloor \frac{q-1}{q-z}\rfloor^t - 1]$.

When $q < 2z$, $1 < z$, and $j \in [1:\lfloor \frac{q-1}{q-z}\rfloor^t - 1]$, let us consider $\mathbf{P}_j$ with row indices generated by $\mathcal{F}^{(j)}_{\mathbf{g}_j} = \{(\mathbf{f}^{(j)}_s, \mathbf{g}_j) \mid s \in [0:q^{m-1}-1]\}$, where $\mathbf{f}^{(j)}_s \in \mathbf{F}_{\mathbf{g}_j}$ and $\mathbf{g}_j = (g^{(j)}_0, g^{(j)}_1, \ldots, g^{(j)}_{t-1})$. For any $\mathbf{v}$ in $\mathbf{P}_0$, we also denote the collection of rows in which $\mathbf{v}$ occurs corresponding to column $(\mathcal{I}, \mathbf{c}) = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, c_1, \ldots, c_{t-1}))$ as $\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}} = \{(\mathbf{f}^{(0)}_s, \mathbf{g}_0) = ((f^{(0)}_{s,0}, f^{(0)}_{s,1}, \ldots, f^{(0)}_{s,m-1}), \mathbf{g}_0) \mid \mathbf{P}((\mathbf{f}^{(0)}_s, \mathbf{g}_0), (\mathcal{I}, \mathbf{c})) = (\mathbf{v}, o(\mathbf{v})), s \in [0:q^{m-1}-1]\}$.
Similarly, suppose that such vector $\mathbf{v}$ occurs in $\mathbf{P}_j$, then the collection of rows in which $\mathbf{v}$ occurs corresponding to column $(\mathcal{I}, \mathbf{c}') = (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c'_0, c'_1, \ldots, c'_{t-1}))$ can be written as $\mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}} = \{(\mathbf{f}^{(j)}_{s'}, \mathbf{g}_j) = ((f^{(j)}_{s',0}, f^{(j)}_{s',1}, \ldots, f^{(j)}_{s',m-1}), \mathbf{g}_j) \mid \mathbf{P}((\mathbf{f}^{(j)}_{s'}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c}')) = (\mathbf{v}, o(\mathbf{v})), s' \in [0:q^{m-1}-1]\}$.
Since $\mathbf{F}_{\mathbf{g}_j}$ is a POA$(m, q, m-1)$ such that $\sum_{r=0}^{m-1} f^{(j)}_{s'r} = \sum_{i=0}^{t-1} g^{(j)}_i (q-z)$ for $s' \in [0:q^{m-1}-1]$, we can define a mapping ϕ from $\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}}$ to $\mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}}$ as

$$\phi((\mathbf{f}^{(0)}_s, \mathbf{g}_0)) = (\mathbf{f}^{(j)}_{s'}, \mathbf{g}_j), (\mathbf{f}^{(0)}_s, \mathbf{g}_0) \in \mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}},\ (\mathbf{f}^{(j)}_{s'}, \mathbf{g}_j) \in \mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}},$$

where $f^{(0)}_{s,i} = f^{(j)}_{s',i}$ for $i \in [0:m-1] \setminus \{\xi_0, \xi_1, \ldots, \xi_{t-1}\}$ and $f^{(0)}_{s,\xi_h} + g^{(j)}_h (q-z) = f^{(j)}_{s',\xi_h}$ for $h \in [0:t-1]$. Given any $h \in [0:t-1]$, let us assume that $f^{(0)}_{s,\xi_h} = c_h + \eta_h$ for $\eta_h \in [1:q-z]$ and $f^{(j)}_{s',\xi_h} = c'_h + \eta'_h$. Based on *Construction 1*, we have $c_h = c'_h - g^{(j)}_h (q-z)$. This implies that $c'_h + \eta'_h = f^{(j)}_{s',\xi_h} = f^{(0)}_{s,\xi_h} + g^{(j)}_h (q-z) = c_h + \eta_h + g^{(j)}_h (q-z) = c'_h - g^{(j)}_h (q-z) + \eta_h + g^{(j)}_h (q-z) = c'_h + \eta_h$, i.e., $\eta_h = \eta'_h$ holds. Hence, given any $(\mathbf{f}^{(0)}_s, \mathbf{g}_0) \in \mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}}$, we can find a unique row $(\mathbf{f}^{(j)}_{s'}, \mathbf{g}_j) \in \mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}}$ such that $\phi((\mathbf{f}^{(0)}_s, \mathbf{g}_0)) = (\mathbf{f}^{(j)}_{s'}, \mathbf{g}_j)$. This means that ϕ is indeed a mapping from $\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}}$ to $\mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}}$. Further based on the definition of ϕ, it can be seen that ϕ is an injective mapping from $\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}}$ to $\mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}}$. Then, we have $|\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}}| \leq |\mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}}|$.

Also, we can define a mapping ϕ' from $\mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}}$ to $\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}}$ as

$$\phi'((\mathbf{f}^{(j)}_{s'}), \mathbf{g}_j) = (\mathbf{f}^{(0)}_{s}, \mathbf{g}_0), \quad (\mathbf{f}^{(j)}_{s'}, \mathbf{g}_j) \in \mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}}, (\mathbf{f}^{(0)}_{s}, \mathbf{g}_0) \in \mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}},$$

where $f^{(j)}_{s',i} = f^{(0)}_{s,i}$ for $i \in [0 : m-1] \setminus \{\xi_0, \xi_1, \ldots, \xi_{t-1}\}$ and $f^{(0)}_{s,\xi_h} + g^{(j)}_h(q-z) = f^{(j)}_{s',\xi_h}$ for $h \in [0 : t-1]$. With the same analysis, we have $|\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}}| \geq |\mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}}|$. Hence, we have $|\mathcal{F}^{(0)}_{\mathcal{I},\mathbf{v}}| = |\mathcal{F}^{(j)}_{\mathcal{I},\mathbf{v}}|$. For any $\mathcal{I}' \neq \mathcal{I}$, with the similar argument above, we also have $|\mathcal{F}^{(0)}_{\mathcal{I}',\mathbf{v}}| = |\mathcal{F}^{(j)}_{\mathcal{I}',\mathbf{v}}|$. Therefore, array $\mathbf{P}_j$ satisfies the statement of (a).

(b). According to the proof of (a), the statement of (b) is clear. ■

Now let us prove *Theorem 1*.

Proof: Let $\mathbf{P}$ be the array generated from *Construction 1*. It can be observed that a vector $\mathbf{v}$ occurring in the column $(\mathcal{I}, \mathbf{c})$ of $\mathbf{P}_j$ does not occur in the column $(\mathcal{I}, \mathbf{c})$ of $\mathbf{P}_{j'}$ for $j' \neq j$ and $j, j' \in [0 : \lfloor \frac{q-1}{q-z} \rfloor^t - 1]$. Let us assume that a vector $\mathbf{v}$ occurring in the column $(\mathcal{I}, \mathbf{c})$ of $\mathbf{P}_j$ occurs in the column $(\mathcal{I}, \mathbf{c})$ of $\mathbf{P}_{j'}$ for $j' \neq j$. Let $\mathbf{g}_j = (g^{(j)}_0, g^{(j)}_1, \ldots, g^{(j)}_{t-1})$, $\mathbf{g}_{j'} = (g^{(j')}_0, g^{(j')}_1, \ldots, g^{(j')}_{t-1})$ and $\mathbf{c} = (c_0, c_1, \ldots, c_{t-1})$ be the vectors defined in *Construction 1*. Without loss of generality, we can assume that $g^{(j)}_i \neq g^{(j')}_i$ for some integer $i \in [0 : t-1]$. Based on *Construction 1*, we have $c_i - g^{(j)}_i(q-z) = c_i - g^{(j')}_i(q-z)$. This is impossible since $g^{(j)}_i(q-z), g^{(j')}_i(q-z) \in [0 : z-1]$. Based on *Proposition 3*, we also have that each vector $\mathbf{v}$ of $\mathbf{P}$ occurs in exactly $\binom{m}{t}\lfloor \frac{q-1}{q-z} \rfloor^t$ columns, and $\mathbf{v}$ occurs the same number of times in each column that contains $\mathbf{v}$, i.e., array $\mathbf{P}$ is a PDA such that the coding gain of each vector entry $(\mathbf{v}, o(\mathbf{v}))$ is $\binom{m}{t}\lfloor \frac{q-1}{q-z} \rfloor^t$. Further based on *Proposition 2*, it can be seen that each column of $\mathbf{P}$ has $\lfloor \frac{q-1}{q-z} \rfloor^t [q^{m-1} - q^{m-t-1}(q-z)^t]$ "$*$"s and the number of non-star entries in $\mathbf{P}$ is $\lfloor \frac{q-1}{q-z} \rfloor^t \binom{m}{t} q^{m-1}(q-z)^t$. This implies that the number of distinct vector entries in $\mathbf{P}$ is $\frac{\lfloor \frac{q-1}{q-z} \rfloor^t \binom{m}{t} q^{m-1}(q-z)^t}{\binom{m}{t}\lfloor \frac{q-1}{q-z} \rfloor^t} = q^{m-1}(q-z)^t$, since the occurences of each vector entry $(\mathbf{v}, o(\mathbf{v}))$ is $\binom{m}{t}\lfloor \frac{q-1}{q-z} \rfloor^t$. Therefore, $\mathbf{P}$ is an $(\binom{m}{t}q^t, \lfloor \frac{q-1}{q-z} \rfloor^t q^{m-1}, \lfloor \frac{q-1}{q-z} \rfloor^t [q^{m-1} - q^{m-t-1}(q-z)^t], q^{m-1}(q-z)^t)$ PDA with a memory ratio of $\frac{M}{N} = 1 - (\frac{q-z}{q})^t$ and a transmission rate of $R = \frac{(q-z)^t}{\lfloor \frac{q-1}{q-z} \rfloor^t}$. ■

Appendix B
Proof of Theorem 2

Proof: Let $\hat{\mathbf{P}} = (\hat{\mathbf{P}}_0 \ \ldots \ \hat{\mathbf{P}}_j \ \ldots \ \hat{\mathbf{P}}_{\lfloor \frac{q-1}{q-z} \rfloor^t - 1})$ denote an array obtained from the transform of *Construction 1*. It can be seen that it is impossible for a vector entry to appear more than once in each column of $\hat{\mathbf{P}}$, since $\hat{\mathbf{P}}_j$ is obtained by selecting some columns of $\mathbf{P}'_j$ and $\mathbf{P}'_j$ is a PDA for $j \in [0 : \lfloor \frac{q-1}{q-z} \rfloor^t - 1]$ (this can be checked with the similar proof of *Proposition 1*). Note that the first $m-1$ coordinates of the row indices of $\hat{\mathbf{P}}_j$ are arranged in the lexicographic order from top to bottom, and $\hat{\mathbf{P}}_j$ $(j \neq 0)$ is obtained by selecting the columns of $\mathbf{P}'_j$ indexed by $(\mathcal{I}, \mathbf{c})$, where $\mathcal{I} \in \binom{[0:m-2]}{t}$ and $\mathbf{c} \in \{(c_0, c_1, \ldots, c_{t-1}) \mid c_i \in [0 : q-1], i \in [0 : t-1]\}$. Assume that a vector entry $\mathbf{e} = (\mathbf{v}, o(\mathbf{v}))$ occurs more than once in some row of $\hat{\mathbf{P}}$. Without loss of generality, let us assume that $\hat{\mathbf{P}}_j((\mathbf{f}^{(j)}_s, \mathbf{g}_j), (\mathcal{I}, \mathbf{c})) = \hat{\mathbf{P}}_{j'}((\mathbf{f}^{(j')}_{s'}, \mathbf{g}_{j'}), (\mathcal{I}', \mathbf{c}')) = \mathbf{e}$ for $j \neq j'$, where

$$\begin{aligned}
(\mathbf{f}^{(j)}_s, \mathbf{g}_j) &= ((f^{(j)}_{s,0}, f^{(j)}_{s,1}, \ldots, f^{(j)}_{s,m-2}, f^{(j)}_{s,m-1}), (g^{(j)}_0, g^{(j)}_1, \ldots, \\
&\qquad g^{(j)}_{t-1}));\\
(\mathbf{f}^{(j')}_{s'}, \mathbf{g}_{j'}) &= ((f^{(j)}_{s,0}, f^{(j)}_{s,1}, \ldots, f^{(j)}_{s,m-2}, f^{(j')}_{s',m-1}), (g^{(j')}_0, g^{(j')}_1, \\
&\qquad \ldots, g^{(j')}_{t-1}));\\
(\mathcal{I}, \mathbf{c}) &= (\{\xi_0, \xi_1, \ldots, \xi_{t-1}\}, (c_0, c_1, \ldots, c_{t-1}));\\
(\mathcal{I}', \mathbf{c}') &= (\{\xi'_0, \xi'_1, \ldots, \xi'_{t-1}\}, (c'_0, c'_1, \ldots, c'_{t-1})).
\end{aligned}$$

Let us consider the following two cases.

Case 1: $|\mathcal{I} \cap \mathcal{I}'| = t$ and $\mathcal{I}, \mathcal{I}' \in \binom{[0:m-2]}{t}$. Assume that the vector entry $\mathbf{e}$ occurs in row $(\mathbf{f}^{(j)}_s, \mathbf{g}_j) = ((f^{(j)}_{s,0}, f^{(j)}_{s,1}, \ldots, f^{(j)}_{s,m-2}, f^{(j)}_{s,m-1}), (g^{(j)}_0, g^{(j)}_1, \ldots, g^{(j)}_{t-1}))$ of $\hat{\mathbf{P}}_j$, then $\mathbf{e}$ must occur in row $(\mathbf{f}^{(j')}_{s'}, \mathbf{g}_{j'}) = ((f^{(j')}_{s',0}, f^{(j')}_{s',1}, \ldots, f^{(j')}_{s',m-2}, f^{(j')}_{s',m-1}), (g^{(j')}_0, g^{(j')}_1, \ldots, g^{(j')}_{t-1}))$ of $\hat{\mathbf{P}}_{j'}$. We claim that $(f^{(j)}_{s,0}, f^{(j)}_{s,1}, \ldots, f^{(j)}_{s,m-2}) \neq (f^{(j')}_{s',0}, f^{(j')}_{s',1}, \ldots, f^{(j')}_{s',m-2})$. This can be verified as follows. Suppose that $\mathbf{e}$ occurs in row $((f^{(0)}_{s'',0}, f^{(0)}_{s'',1}, \ldots, f^{(0)}_{s'',m-2}, f^{(0)}_{s'',m-1}), \mathbf{g}_0)$ of $\mathbf{P}'_0$. According to $\mathbf{g}_j \neq \mathbf{g}_{j'}$, it can be seen that the hamming distance between two vectors $\mathbf{g}_j$ and $\mathbf{g}_{j'}$ is at least one. Without loss of generality, assume that there exists an integer $\alpha \in [0 : t-1]$ such that $g^{(j)}_\alpha \neq g^{(j')}_\alpha$. Based on the process of the transform and the proof of *Proposition 3*, we obtain $(f^{(j)}_{s,0}, f^{(j)}_{s,1}, \ldots, f^{(j)}_{s,m-2}) = (f^{(0)}_{s'',0}, \ldots, f^{(0)}_{s'',\xi_0} + g^{(j)}_0(q-z), \cdots, f^{(0)}_{s'',\xi_\alpha} + g^{(j)}_\alpha(q-z), \ldots, f^{(0)}_{s'',\xi_{t-1}} + g^{(j)}_{t-1}(q-z), \ldots, f^{(0)}_{s'',m-2})$ and $(f^{(j')}_{s,0}, f^{(j')}_{s,1}, \ldots, f^{(j')}_{s,m-2}) = (f^{(0)}_{s'',0}, \ldots, f^{(0)}_{s'',\xi_0} + g^{(j)}_0(q-z), \ldots, f^{(0)}_{s'',\xi_\alpha} + g^{(j')}_\alpha(q-z), \ldots, f^{(0)}_{s'',\xi_{t-1}} + g^{(j)}_{t-1}(q-z), \ldots, f^{(0)}_{s'',m-2})$. Then, we have $f^{(0)}_{s'',\xi_\alpha} + g^{(j)}_\alpha(q-z) \neq f^{(0)}_{s'',\xi_\alpha} + g^{(j')}_\alpha(q-z)$ since $g^{(j)}_\alpha(q-z), g^{(j')}_\alpha(q-z) \in [0 : z-1]$. This implies that $(f^{(j)}_{s,0}, f^{(j)}_{s,1}, \ldots, f^{(j)}_{s,m-2}) \neq (f^{(j')}_{s',0}, f^{(j')}_{s',1}, \ldots, f^{(j')}_{s',m-2})$.

Case 2: $|\mathcal{I} \cap \mathcal{I}'| < t$ and $\mathcal{I}' \in \binom{[0:m-2]}{t}$. There must exist two distinct integers, say α, $\alpha' \in [0 : m-1]$, satisfying $\alpha \in \mathcal{I}, \alpha \notin \mathcal{I}'$ and $\alpha' \in \mathcal{I}', \alpha' \notin \mathcal{I}$. Without loss of generality, let us assume that $\alpha = \xi_0$. Based on the construction of $\hat{\mathbf{P}}_j$ and $\hat{\mathbf{P}}_{j'}$, we obtain $f^{(j)}_{s,\xi_0} = c_0 - g^{(j)}_0(q-z)$. This implies that $\hat{\mathbf{P}}_j((\mathbf{f}^{(j)}_s, \mathbf{g}_j), (\mathcal{I}, \mathbf{c})) = *$ since $f^{(j)}_{s,\xi_0} = c_0 - g^{(j)}_0(q-z) \in \{c_0, c_0-1, \ldots, c_0-(z-1)\}$, which contradicts the hypothesis.

Based on the above argument, it can be seen that C3-(a) of *Definition 1* holds. It remains to show that C3-(b) of *Definition 1* holds for $\hat{\mathbf{P}}$. Note that $\hat{\mathbf{P}}_j$ is a PDA for $j \in [1 : \lfloor \frac{q-1}{q-z} \rfloor^t - 1]$. To do this, one just needs to prove if $\hat{\mathbf{P}}_j((\mathbf{f}^{(j)}_s, \mathbf{g}_j), (\mathcal{I}, \mathbf{c})) = \hat{\mathbf{P}}_{j'}((\mathbf{f}^{(j')}_{s'}, \mathbf{g}_{j'}), (\mathcal{I}', \mathbf{c}')) = \mathbf{e}$ for $j \neq j'$ and $\sum_{i=0}^{t-1} g^{(j)}_i \geq \sum_{i=0}^{t-1} g^{(j')}_i$, then $\hat{\mathbf{P}}_j((\hat{\mathbf{f}}^{(j')}_{s'}, \mathbf{g}_j), (\mathcal{I}, \mathbf{c})) = \hat{\mathbf{P}}_{j'}((\hat{\mathbf{f}}^{(j)}_{s}, \mathbf{g}_{j'}), (\mathcal{I}', \mathbf{c}')) = *$, where

$$\begin{aligned}
(\mathbf{f}^{(j)}_s, \mathbf{g}_j) &= ((f^{(j)}_{s,0}, f^{(j)}_{s,1}, \ldots, f^{(j)}_{s,m-2}, f^{(j)}_{s,m-1}), (g^{(j)}_0, g^{(j)}_1, \ldots, \\
&\qquad g^{(j)}_{t-1}));\\
(\mathbf{f}^{(j')}_{s'}, \mathbf{g}_{j'}) &= ((f^{(j')}_{s',0}, f^{(j')}_{s',1}, \ldots, f^{(j')}_{s',m-2}, f^{(j')}_{s',m-1}), (g^{(j')}_0, g^{(j')}_1, \\
&\qquad \ldots, g^{(j')}_{t-1}));
\end{aligned}$$

$$(\mathcal{I},\mathbf{c}) = (\{\xi_0,\xi_1,\ldots,\xi_{t-1}\},(c_0,c_1,\ldots,c_{t-1}));$$
$$(\mathcal{I}',\mathbf{c}') = (\{\xi'_0,\xi'_1,\ldots,\xi'_{t-1}\},(c'_0,c'_1,\ldots,c'_{t-1}));$$
$$\hat{\mathbf{f}}_{s'}^{(j')} = (f_{s',0}^{(j')},f_{s',1}^{(j')},\ldots,f_{s',m-2}^{(j')},f_{s',m-1}^{(j')}+x(q-z));$$
$$\hat{\mathbf{f}}_{s}^{(j)} = (f_{s,0}^{(j)},f_{s,1}^{(j)},\ldots,f_{s,m-2}^{(j)},f_{s,m-1}^{(j)}-x(q-z));$$
$$x = \sum_{i=0}^{t-1} g_i^{(j)} - \sum_{i=0}^{t-1} g_i^{(j')}.$$

Let us consider the following two cases.

Case 3: $\mathcal{I},\mathcal{I}' \in \binom{[0:m-2]}{t}$. Since $(\hat{\mathbf{P}}_j;\hat{\mathbf{P}}_{j'})$ is a PDA, we have $\hat{\mathbf{P}}_{j'}((\mathbf{f}_{s'}^{(j')},\mathbf{g}_{j'}),(\mathcal{I},\mathbf{c})) = \hat{\mathbf{P}}_j((\mathbf{f}_s^{(j)},\mathbf{g}_j),(\mathcal{I}',\mathbf{c}')) = *$. This implies that $\hat{\mathbf{P}}_j((\hat{\mathbf{f}}_{s'}^{(j')},\mathbf{g}_j),(\mathcal{I},\mathbf{c})) = \hat{\mathbf{P}}_{j'}((\hat{\mathbf{f}}_s^{(j)},\mathbf{g}_{j'}),(\mathcal{I}',\mathbf{c}')) = *$ due to the entry rule in *Construction 1*.

Case 4: $\mathcal{I} \in \binom{[0:m-2]}{t}$ and $\mathcal{I}' \notin \binom{[0:m-2]}{t}$. In this case, one just needs to consider $j'=0$, i.e., one just needs to prove if $\hat{\mathbf{P}}_j((\mathbf{f}_s^{(j)},\mathbf{g}_j),(\mathcal{I},\mathbf{c})) = \hat{\mathbf{P}}_0((\mathbf{f}_{s'}^{(0)},\mathbf{g}_0),(\mathcal{I}',\mathbf{c}')) = \mathbf{e}$ for $j \neq 0$, then $\hat{\mathbf{P}}_j((\hat{\mathbf{f}}_{s'}^{(0)},\mathbf{g}_j),(\mathcal{I},\mathbf{c})) = \hat{\mathbf{P}}_0((\hat{\mathbf{f}}_s^{(j)},\mathbf{g}_0),(\mathcal{I}',\mathbf{c}')) = *$, where $\mathbf{g}_0 = (g_0^{(0)},g_1^{(0)},\ldots,g_{t-1}^{(0)}) = (0,0,\ldots,0)$. With the similar argument of Case 3, it can be seen that $\hat{\mathbf{P}}_j((\hat{\mathbf{f}}_{s'}^{(0)},\mathbf{g}_j),(\mathcal{I},\mathbf{c})) = *$ holds. Next we show that $\hat{\mathbf{P}}_0((\hat{\mathbf{f}}_s^{(j)},\mathbf{g}_0),(\mathcal{I}',\mathbf{c}')) = *$ also holds. If $|\mathcal{I}\cap\mathcal{I}'| < t-1$, based on entry rule of $\hat{\mathbf{P}}$ and *Construction 1*, there must exist an integer $\xi_\alpha \in \mathcal{I}'$ such that $f_{s',\xi_\alpha}^{(0)} = c'_\alpha - g_\alpha^{(0)}(q-z)$. This implies that $\hat{\mathbf{P}}_0((\hat{\mathbf{f}}_s^{(j)},\mathbf{g}_0),(\mathcal{I}',\mathbf{c}')) = *$ since $f_{s',\xi_\alpha}^{(0)} = c'_\alpha$. If $|\mathcal{I}\cap\mathcal{I}'| = t-1$, without loss of generality, let us assume that $\xi_{\alpha_0} = \xi'_{\beta_0}, \xi_{\alpha_1} = \xi'_{\beta_1}, \ldots, \xi_{\alpha_{t-2}} = \xi'_{\beta_{t-2}}$, where $\{\xi_{\alpha_0},\xi_{\alpha_1},\ldots,\xi_{\alpha_{t-2}}\} \subseteq \mathcal{I}$ and $\{\xi'_{\beta_0},\xi'_{\beta_1},\ldots,\xi'_{\beta_{t-2}}\} \subseteq \mathcal{I}'$. Based on *Construction 1*, we have

$$c_{\alpha_0} - g_{\alpha_0}^{(j)}(q-z) = c'_{\beta_0} - g_{\beta_0}^{(0)}(q-z) = c'_{\beta_0};$$
$$c_{\alpha_1} - g_{\alpha_1}^{(j)}(q-z) = c'_{\beta_1} - g_{\beta_1}^{(0)}(q-z) = c'_{\beta_1};$$
$$\vdots$$
$$c_{\alpha_{t-2}} - g_{\alpha_{t-2}}^{(j)}(q-z) = c'_{\beta_{t-2}} - g_{\beta_{t-2}}^{(0)}(q-z) = c'_{\beta_{t-2}};$$
$$f_{s,m-1}^{(j)} = c'_{t-1} - g_{t-1}^{(0)}(q-z) = c'_{t-1}.$$

Suppose that $f_{s,\xi'_{\beta_i}}^{(j)} = c'_{\beta_i} + \gamma$ for $\gamma \in [1:q-z]$ and $i \in [0:t-2]$, i.e., $\hat{\mathbf{P}}_0((\hat{\mathbf{f}}_s^{(j)},\mathbf{g}_0),(\mathcal{I}',\mathbf{c}')) \neq *$. Then, we obtain

$$c_{\alpha_0} - g_{\alpha_0}^{(j)}(q-z) = c'_{\beta_0} = f_{s,\xi'_{\beta_0}}^{(j)} - \gamma;$$
$$\vdots$$
$$c_{\alpha_{t-3}} - g_{\alpha_{t-3}}^{(j)}(q-z) = f_{s,\xi'_{\beta_{t-3}}}^{(j)} - \gamma;$$
$$c_{\alpha_{t-2}} - g_{\alpha_{t-2}}^{(j)}(q-z) = f_{s,\xi'_{\beta_{t-2}}}^{(j)} - \gamma.$$

This implies that $c_{\alpha_i} + \gamma - g_{\alpha_i}^{(j)}(q-z) = f_{s,\xi'_{\beta_i}}^{(j)} = f_{s,\xi_{\alpha_i}}^{(j)}$ for $i \in [0:t-2]$. If there exists an integer $i \in [0:t-2]$ such that $g_{\alpha_i}^{(j)} \neq 0$, we have $\hat{\mathbf{P}}_j((\mathbf{f}_s^{(j)},\mathbf{g}_j),(\mathcal{I},\mathbf{c})) = *$ due to $g_{\alpha_i}^{(j)}(q-z) \in [q-z:z-1]$ and $q < 2z$, which contradicts $\hat{\mathbf{P}}_j((\mathbf{f}_s^{(j)},\mathbf{g}_j),(\mathcal{I},\mathbf{c})) = \mathbf{e}$. So $\gamma \notin [1:q-z]$, i.e., $\hat{\mathbf{P}}_0((\hat{\mathbf{f}}_s^{(j)},\mathbf{g}_0),(\mathcal{I}',\mathbf{c}')) = *$. If $g_{\alpha_i}^{(j)} = 0$ for any $i \in [0:t-2]$, it can be seen that $g_{t-1}^{(j)} \neq 0$ since $\mathbf{g}_j \neq \mathbf{g}_0 = (0,0,\ldots,0)$. Note that $f_{s,m-1}^{(j)} = c'_{t-1}$. This implies that $f_{s,m-1}^{(j)} - x(q-z) \in \{c'_{t-1}, c'_{t-1}-1,\ldots,c'_{t-1}-(z-1)\}$, where $x = \sum_{i=0}^{t-1} g_i^{(j)} - \sum_{i=0}^{t-1} g_i^{(0)} = g_{t-1}^{(j)}$. Hence, we also have $\hat{\mathbf{P}}_0((\hat{\mathbf{f}}_s^{(j)},\mathbf{g}_0),(\mathcal{I}',\mathbf{c}')) = *$. Therefore, C3-(b) of *Definition 1* holds.

Since $\hat{\mathbf{P}} = (\hat{\mathbf{P}}_0\ \ldots\ \hat{\mathbf{P}}_j\ \ldots\ \hat{\mathbf{P}}_{\lfloor\frac{q-1}{q-z}\rfloor^t-1})$ and $\hat{\mathbf{P}}_j$ is a PDA for $j \in [0:\lfloor\frac{q-1}{q-z}\rfloor^t-1]$, it can be seen that the number of "$*$"s in each column of $\hat{\mathbf{P}}$ is $q^{m-1}-q^{m-t-1}(q-z)^t$. Further based on *Proposition 3*, it can be seen that the number distinct entries in $\hat{\mathbf{P}}$ is $q^{m-1}(q-z)^t$. Therefore, array $\hat{\mathbf{P}}$ is an $(((\binom{m-1}{t}\lfloor\frac{q-1}{q-z}\rfloor^t + \binom{m}{t} - \binom{m-1}{t})q^t,\ q^{m-1},\ q^{m-1}-q^{m-t-1}(q-z)^t,\ q^{m-1}(q-z)^t)$ PDA with a memory ratio of $\frac{M}{N} = 1-(\frac{q-z}{q})^t$ and a transmission rate of $R = (q-z)^t$. ■

References

[1] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[2] M. M. Amiri and D. Gündüz, "Fundamental limits of coded caching: Improved delivery rate-cache capacity tradeoff," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 806–815, Feb. 2017.

[3] K. Zhang and C. Tian, "Fundamental limits of coded caching: From uncoded prefetching to coded prefetching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1153–1164, Jun. 2018.

[4] H. Ghasemi and A. Ramamoorthy, "Improved lower bounds for coded caching," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Hong Kong, Jun. 2015, pp. 1696–1700.

[5] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "Characterizing the rate-memory tradeoff in cache networks within a factor of 2," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 647–663, Jan. 2019.

[6] K. Wan, D. Tuninetti, and P. Piantanida, "On the optimality of uncoded cache placement," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Cambridge, U.K., Sep. 2016, pp. 161–165.

[7] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "The exact rate-memory tradeoff for caching with uncoded prefetching," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1281–1296, Feb. 2018.

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

[9] F. Arbabjolfaei and Y.-H. Kim, "Fundamentals of index coding," *Found. Trends Commun. Inf. Theory*, vol. 14, nos. 3–4, pp. 163–346, 2018.

[10] J. Gómez-Vilardebó, "Fundamental limits of caching: Improved rate memory trade-off with coded prefetching," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4488–4497, Oct. 2018.

[11] C. Tian and J. Chen, "Caching and delivery via interference elimination," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1548–1560, Mar. 2018.

[12] M. Ji, K. Shanmugam, G. Vettigli, J. Llorca, A. M. Tulino, and G. Caire, "An efficient multiple-groupcast coded multicasting scheme for finite fractional caching," in *Proc. IEEE ICC*, Jun. 2015, pp. 3801–3806.

[13] S. Jin, Y. Cui, H. Liu, and G. Caire, "Structural properties of uncoded placement optimization for coded delivery," 2017, *arXiv:1707.07146*.

[14] K. Shanmugam, M. Ji, A. M. Tulino, J. Llorca, and A. G. Dimakis, "Finite length analysis of caching-aided coded multicasting," *IEEE Trans. Inf. Theory*, vol. 62, no. 10, pp. 5524–5537, Oct. 2016.

[15] S. Agrawal, K. V. S. Sree, and P. Krishnan, "Coded caching based on combinatorial designs," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 1227–1231.

[16] H. H. S. Chittoor, M. Bhavana, and P. Krishnan, "Coded caching via projective geometry: A new low subpacketization scheme," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 682–686.

[17] H. H. S. Chittoor, P. Krishnan, K. V. S. Sree, and B. Mamillapalli, "Subexponential and linear subpacketization coded caching via projective geometry," *IEEE Trans. Inf. Theory*, vol. 67, no. 9, pp. 6193–6222, Sep. 2021.

[18] C. Shangguan, Y. Zhang, and G. Ge, "Centralized coded caching schemes: A hypergraph theoretical approach," *IEEE Trans. Inf. Theory*, vol. 64, no. 8, pp. 5755–5766, Aug. 2018.

[19] K. Shanmugam, A. M. Tulino, and A. G. Dimakis, "Coded caching with linear subpacketization is possible using Ruzsa-Szeméredi graphs," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 1237–1241.

[20] K. Shanmugam, A. G. Dimakis, J. Llorca, and A. M. Tulino, "A unified Ruzsa-Szemerédi framework for finite-length coded caching," in *Proc. 51st ACSSC*, Pacific Grove, CA, Oct. 2017, pp. 631–635.
[21] Q. Yan, M. Cheng, X. Tang, and Q. Chen, "On the placement delivery array design for centralized coded caching scheme," *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 5821–5833, Sep. 2017.
[22] Q. Yan, X. Tang, Q. Chen, and M. Cheng, "Placement delivery array design through strong edge coloring of bipartite graphs," *IEEE Commun. Lett.*, vol. 22, no. 2, pp. 236–239, Feb. 2018.
[23] L. Tang and A. Ramamoorthy, "Coded caching schemes with reduced subpacketization from linear block codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 4, pp. 3099–3120, Apr. 2018.
[24] M. Cheng, J. Jiang, X. Tang, and Q. Yan, "Some variant of known coded caching schemes with good performance," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1370–1377, Mar. 2020.
[25] M. Cheng, J. Jiang, Q. Yan, and X. Tang, "Constructions of coded caching schemes with flexible memory size," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4166–4176, Jun. 2019.
[26] M. Cheng, J. Wang, X. Zhong, and Q. Wang, "A framework of constructing placement delivery arrays for centralized coded caching," *IEEE Trans. Inf. Theory*, vol. 67, no. 11, pp. 7121–7131, Nov. 2021.
[27] Z. Mingming, C. Minquan, W. Jinyu, Z. Xi, and Y. Chen, "Improving placement delivery array coded caching schemes with coded placement," *IEEE Access*, vol. 8, pp. 217456–217462, 2020.
[28] M. Cheng, J. Jiang, Q. Wang, and Y. Yao, "A generalized grouping scheme in coded caching," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3422–3430, May 2019.
[29] X. Zhong, M. Cheng, and J. Jiang, "Placement delivery array based on concatenating construction," *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1216–1220, Jun. 2020.
[30] J. Michel and Q. Wang, "Placement delivery arrays from combinations of strong edge colorings," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 5953–5964, Oct. 2020.
[31] D. R. Stinson, *Combinatorial Designs: Construction and Analysis*. New York, NY, USA: Springer, 2003.

Xianzhang Wu received the B.S. degree in mathematics and applied mathematics from Minjiang University, Fuzhou, China, in 2014, and the M.S. degree in applied mathematics from Fuzhou University, Fuzhou, in 2018. He is currently pursuing the Ph.D. degree with the School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen, China. His research interests include combinatorics, graph theory, and caching networks and their interactions.

Minquan Cheng received the Ph.D. degree from the Department of Social Systems and Management, Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Ibaraki, Japan, in 2012. Then, he joined Guangxi Normal University, Guilin, Guangxi, China, where he is currently a Full Professor with the School of Computer Science and Information Technology. His research interests include combinatorics, coding theory, and cryptography and their interactions.

Congduan Li (Member, IEEE) received the B.S. degree in electrical engineering from the University of Science and Technology Beijing, China, in 2008, the M.S. degree in electrical engineering degree from Northern Arizona University, AZ, USA, in 2011, and the Ph.D. degree in electrical engineering from Drexel University, PA, USA, in 2015, respectively. From October 2015 to August 2018, he was a Post-Doctoral Research Fellow with the Institute of Network Coding, The Chinese University of Hong Kong, and the Department of Computer Science, City University of Hong Kong. He is currently an Associate Professor with the School of Electronics and Communication Engineering, Sun Yat-sen University, China. His research interests lie in the broad areas related with networks, such as coding, security, wireless, storage, and caching.

Li Chen (Senior Member, IEEE) received the B.Sc. degree in applied physics from Jinan University, China, in 2003, and the M.Sc. degree in communications and signal processing and the Ph.D. degree in communications engineering from Newcastle University, U.K., in 2004 and 2008, respectively. From 2007 to 2010, he was a Research Associate with Newcastle University. In 2010, he returned China as a Lecturer at the School of Information Science and Technology, Sun Yat-sen University, Guangzhou. From 2011 to 2012, he was a Visiting Researcher with the Institute of Network Coding, The Chinese University of Hong Kong. From 2011 and 2016, he was an Associate Professor and a Professor of the University. Since 2013, he has been the Associate Head of the Department of Electronic and Communication Engineering (ECE). From July 2015 to October 2015, he was a Visitor at the Institute of Communications Engineering, Ulm University, Germany. From October 2015 to June 2016, he was a Visiting Associate Professor at the Department of Electrical Engineering, University of Notre Dame, USA. From 2017 to 2020, he was the Deputy Dean of the School of Electronics and Communication Engineering. His research interests include information theory, error-correction codes, and data communications. He is a Senior Member of the Chinese Institute of Electronics (CIE). He is a member of the IEEE Information Theory Society Board of Governors and its External Nomination Committee and the Chair of its Conference Committee. He is also the Chair of the IEEE Information Theory Society Guangzhou Chapter. He is an Associate Editor of IEEE TRANSACTIONS ON COMMUNICATIONS. He has been organizing several international conferences and workshops, including the 2018 IEEE Information Theory Workshop (ITW) at Guangzhou and the 2022 IEEE East Asian School of Information Theory (EASIT) at Shenzhen, for which he is the General Co-Chair. He is also the TPC Co-Chair of the 2022 IEEE/CIC International Conference on Communications in China (ICCC) at Foshan. He likes reading and photography.