On Zero Skip-Cost Generalized Fractional-Repetition Codes from Covering Designs

Wenjun Yu*, Bo-Jun Yuan†, and Moshe Schwartz‡*
* School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer Sheva 8410501, Israel
† School of Science, Zhejiang University of Science and Technology, Hangzhou 310023, China
‡ Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, L8S 4K1, Canada
Emails: wenjun@post.bgu.ac.il, ybjmath@163.com, schwartz.moshe@mcmaster.ca

2026 IEEE International Symposium on Information Theory (ISIT) | 979-8-3315-8004-9/26/$31.00 ©2026 IEEE | DOI: 10.1109/ISIT62367.2026.11653703

***Abstract*—We study generalized fractional repetition codes that have zero skip cost, and which are based on covering designs. We show that a zero skip cost is always attainable, perhaps at a price of an expansion factor compared with the optimal size of fractional repetition codes based on Steiner systems. We provide three constructions, as well as show non-constructively, that no expansion is needed for all codes based on sufficiently large covering systems.**

I. Introduction

Distributed storage systems (DSS) underpin large-scale data analytics, cloud services, and archival platforms. Because node failures are routine, ensuring reliability while maintaining fast and resource-efficient repair has become a central design challenge. This has led to the development of coding schemes tailored for distributed settings, among which fractional repetition (FR) codes [2] have played a particularly influential role. The FR code is embedded within the DRESS (Distributed Replication based Exact Simple Storage) architecture [2]. In this scheme, user information is first encoded using an outer MDS code. The coded information is then coded using an inner FR code, which distributes copies of the coded information across several nodes. These form a class of exact minimum-bandwidth regenerating (MBR) codes. FR codes offer low-complexity uncoded repair of failed nodes, where a failed node can be reconstructed purely through symbol transfer, without any computation at helper nodes.

Since constructions for MDS codes with flexible parameters are known, considerable effort has been devoted to the construction of FR codes. A common strategy for the latter is to base the construction on combinatorial designs (e.g., [7], [2], [14], [17], [6], [16]). Among these, using Steiner systems is perhaps the most popular. However, since only few constructions of Steiner systems are known, and their parameters are quite restrictive, it was recently suggested to consider other designs [18], [15] resulting in generalized fractional-repetition (GFR) codes. These developments retain the simplicity of uncoded repair while supporting a much wider range of system parameters.

When a node fails (node erasure), a small number of helper nodes are contacted. These nodes transfer copies of parts of the erased data, thereby recovering all of the erased node. A key practical aspects of repair performance in storage systems is the number of helper nodes. Practical systems typically enforce limits on the number of helper nodes contacted during repair, commonly referred to as *locality*.

At the same time, repair performance is further affected by the mechanism through which helper nodes retrieve the required symbols from local storage. In many real deployments, including HDD based storage, sequential reads are dramatically faster than random reads, and even slight misalignment in symbol positions can lead to significant seek overhead during repair. Thus, the order in which symbols are stored becomes crucial [11]. When the required symbols are scattered across different positions, the helper experiences additional disk seeks. This behavior, captured by the *skip cost*, directly influences repair time, bandwidth consumption, and overall system efficiency. To illustrate, consider the simple system represented by the array

$$\mathcal{A} = \begin{pmatrix} 1 & 2 & 1 & 2 & 1 & 1 \\ 2 & 3 & 3 & 4 & 3 & 2 \\ 3 & 4 & 4 & 5 & 5 & 4 \\ 5 & 6 & 5 & 6 & 6 & 6 \end{pmatrix}, \tag{1}$$

where each column represents a node and each row represents a coded symbol index within that node. Suppose the system enforces a locality parameter $\ell = 2$, meaning any failed node must be repaired using at most two helper nodes. If Node 1:$(1,2,3,5)$ fails, one valid repair option is to contact helpers Node 3 and Node 6 such that Node 3 contributes the symbols $\{3,5\}$, while Node 6 contributes $\{1,2\}$. Because the two requested symbols in Node 3's storage vector are separated by one irrelevant symbol that needs to be skipped, and two requested symbols in Node 6's storage vector are sequential, this repair incurs a skip cost of 1. Alternatively, contacting helpers Node 5 and Node 6 results in both helpers supplying consecutive symbols $\{3,5\}$ and $\{1,2\}$ in their stored vector, respectively. This option has a zero skip cost.

Several works study coding schemes for storage, that have a small skip cost [1], [5], [13]. In particular, Chee et al. [1] defined the skip cost of the FR code, with zero skip cost being the most desirable. They went on to construct zero skip-cost FR codes based on Steiner quadruple systems. Apart from this, no other non-trivial FR codes with zero skip cost are known.

This motivated the study of zero skip cost FR and GFR codes, in which each helper reads the required repair symbols in strictly consecutive positions. Existing constructions, however, rely almost entirely on Steiner systems, and thus inherit

The research of Bo-Jun Yuan was supported by the National Natural Science Foundation of China under Grant 12201559.

their severe restrictions. When Steiner systems are unavailable, which is the norm for most parameter choices, these designs cannot be applied. To overcome this limitation, this paper develops a general and flexible framework for constructing zero skip-cost generalized fractional repetition codes based on covering designs, which we call covering fractional repetition (CFR) codes. Covering designs are known to exist for all parameters while asymptotically approaching those of Steiner systems, as shown by Rödl [9].

The main contributions of this paper are as follows: We construct CFR codes with zero skip cost. The zero skip cost, however, comes at a price of potentially having more nodes than an FR code based on a Steiner system of the same parameters (if it were to exist in those parameters at all). We call the ratio of obtained nodes to the (mostly unattainable) number of nodes in a Steiner system, the expansion factor. We show a trivial construction with asymptotic expansion factor 2, and then a second construction with asymptotic expansion factor 1. However, in finite cases, the expansion factor of the second construction might be higher than that of the first. We then provide a third more efficient construction for finite cases, improving upon the first two in many cases. Finally, we show using a probabilistic non-constructive argument, that for all sufficiently-large covering designs, no expansion is required whatsoever in order to attain zero skip cost.

The paper is organized as follows. In Section II we give the notation and rigorous definitions used throughout the paper. Our constructions of CFR codes are described in Section III. Due to space limitations, some material has been omitted. The complete version may be found in [12].

II. Preliminaries

For integers $i < j$, we define $[i, j] \triangleq \{i, i+1, \dots, j\}$, and for $v \in \mathbb{N}$, we define $[v] \triangleq [1, v]$. Given a set S, we denote by $\binom{S}{k}$ the set of all k-subsets of S, namely, $\binom{S}{k} \triangleq \{S' \subseteq S : |S'| = k\}$. For any two positive integers, a and b, we let $a \,\overline{\mathrm{mod}}\, b$ denote the unique integer $r \in [b]$ such that $r \equiv a \pmod{b}$. Thus, $a = (\lceil a/b \rceil - 1)b + (a \,\overline{\mathrm{mod}}\, b)$. We shall also use the notation $\{\!\{\}\!\}$ to denote multisets, i.e., sets with possibly repeated elements.

Throughout this paper we work over the alphabet $[v]$, and we consider the distributed storage system based on a $k \times N$ array $\mathcal{A} \in [v]^{k \times N}$ with form $\mathcal{A} = (\boldsymbol{a}^{(1)}, \dots, \boldsymbol{a}^{(N)})$, where $\boldsymbol{a}^{(i)} \in [v]^k$ is column vector of length k. In this framework, each column of $\mathcal{A}$ represents a node. There are v symbols spread among the N nodes such that the i-th node stores k symbols in an ordered sequence $\boldsymbol{a}^{(i)}$.

When some node j fails, the system repairs it by transferring symbols from other nodes. We consider *repair by transfer*, also called *uncoded repair* [8]. The repair process calls upon a set of nodes, called *helper nodes*, $\mathcal{H}^{(j)} \subseteq [N] \setminus \{j\}$, which depends on j. Each node $h \in \mathcal{H}^{(j)}$ transmits a subset of its entries, $R^{(h)}$, back to node j such that the original entries of j are all received, namely, $\bigcup_{h \in \mathcal{H}^{(j)}} R^{(h)}$ is exactly the set of entries in $\boldsymbol{a}^{(j)}$. Since we would like to minimize the bandwidth required for repairing, we further assume no element is transmitted twice, which means that $R^{(h)} \cap R^{(h')} = \emptyset$, for all $h, h' \in \mathcal{H}^{(j)}$, $h \neq h'$. Another standard assumption we make is that the repair process has *locality* ℓ, which means no more than ℓ helper nodes participate in the repair, i.e., $|\mathcal{H}^{(j)}| \leqslant \ell$.

TABLE I
Repairing $\mathcal{A}$ from (1) with zero skip cost

Column	Repair Process	Skip Cost
1	$\{1,2\}$ from $\boldsymbol{a}^{(6)}$, $\{3,5\}$ from $\boldsymbol{a}^{(5)}$	0
2	$\{2,3\}$ from $\boldsymbol{a}^{(1)}$, $\{4,6\}$ from $\boldsymbol{a}^{(6)}$	0
3	$\{1,3\}$ from $\boldsymbol{a}^{(5)}$, $\{4,5\}$ from $\boldsymbol{a}^{(4)}$	0
4	$\{2,4\}$ from $\boldsymbol{a}^{(6)}$, $\{5,6\}$ from $\boldsymbol{a}^{(5)}$	0
5	$\{1,3\}$ from $\boldsymbol{a}^{(3)}$, $\{5,6\}$ from $\boldsymbol{a}^{(4)}$	0
6	$\{1,2\}$ from $\boldsymbol{a}^{(1)}$, $\{4,6\}$ from $\boldsymbol{a}^{(2)}$	0

At this point we would like to consider the *skip cost* of the repair process. The concept was first introduced by [1] as a measure of the repair-process performance.

Definition 1. *Assume the distributed-storage code $\mathcal{A}$, as defined above, with locality ℓ. Further assume that during a repair process, node $h \in [N]$ transmits a subset of its contents, $R^{(h)}$, where $\boldsymbol{a}^{(h)} = (a_1^{(h)}, a_2^{(h)}, \dots, a_k^{(h)})$, and $R^{(h)} = \{a_{i_1}^{(h)}, a_{i_2}^{(h)}, \dots, a_{i_t}^{(h)}\}$, for some $1 \leqslant i_1 < \dots < i_t \leqslant k$. the* skip cost *for node h is defined in this case as*

$$\mathsf{cost}(R^{(h)} \mid \boldsymbol{a}^{(h)}) \triangleq \sum_{j=1}^{t-1} (i_{j+1} - i_j - 1) = i_t - i_1 - (t-1).$$

We naturally extend this to the skip cost of repairing node j,

$$\mathsf{cost}(\boldsymbol{a}^{(j)}) \triangleq \min_{\mathcal{H}^{(j)}} \sum_{h \in \mathcal{H}^{(j)}} \mathsf{cost}(R^{(h)} \mid \boldsymbol{a}^{(h)}),$$

where the minimization is over all ways of finding a set of helper nodes $\mathcal{H}^{(j)}$, $|\mathcal{H}^{(j)}| \leqslant \ell$, and assigning each $h \in \mathcal{H}^{(j)}$ a transmission $R^{(h)}$. The skip cost of the entire array is

$$\mathsf{cost}(\mathcal{A}) \triangleq \max_{j \in [N]} \mathsf{cost}(\boldsymbol{a}^{(j)}).$$

Example 1. *Consider $\mathcal{A}$ from* (1)*. Then $\mathsf{cost}(\mathcal{A}) = 0$ as Table I shows.*

To better understand the considerations in constructing $\mathcal{A}$, we recall the framework of fractional repetition codes. El Rouayheb and Ramchandran [2] introduced the concept of a DRESS (Distributed Replication-based Exact Simple Storage) code, which consists of the concatenation of an outer MDS code and an inner fractional repetition (FR) code. In their setting, assume that k, N, v, ρ are positive integers satisfying $kN = v\rho$. An (N, k, ρ)-FR code is a collection of N k-subsets of $[v]$, $B_1, B_2, \dots, B_N \in \binom{[v]}{k}$, such that each symbol of $[v]$ appears in *exactly* ρ subsets. For convenience, a $k \times N$ matrix $\mathcal{A}$ is defined, whose i-th column, $\boldsymbol{a}^{(i)}$, contains the elements of B_i in some arbitrary order.

Several studies have suggested ways of constructing the blocks $B_1, \dots, B_n$ which define the FR code. Most of these revolve around the use block designs, as these naturally contain points and blocks as part of their definition. In some cases the blocks of the system are used as is, while in some the transpose design is used, where the role of points and blocks is reversed. As examples we mention the use of Steiner systems in [7], [2], [14], [1], t-(v,k,λ) designs in [17], and affine resolvable designs [6], [7], [16].

The popularity of Steiner systems is shadowed by their relative rarity. The strict divisibility conditions required prohibit many parameters, and the few constructions known provide only Steiner systems with small parameters. However, in [18], [15], more flexible codes were suggested, called *generalized fractional repetition (GFR) codes*, paving the way to using other designs, such as covering designs, which are the focus of this paper. To describe these codes, we first recall the definition of covering designs.

Definition 2. *Let $t \leqslant k \leqslant v$ be positive integers. A (t,k,v)* covering design *is the pair $(X,\mathcal{B})$, where $X=[v]$ (the* points*) and a multiset $\mathcal{B}=\{\!\{B_1,\dots,B_N\}\!\} \subseteq \binom{X}{k}$ (the* blocks*), such that any t-subset of points $T \in \binom{X}{t}$ is contained in at least one block.*

This definition generalizes the notion of (t,k,v) Steiner systems, which require that any t-subset is contained in *exactly one* block. Assume $(X,\mathcal{B})$ is a (t,k,v) covering design. Let

$$r_s \triangleq \frac{\binom{v-s}{t-s}}{\binom{k-s}{t-s}}, \tag{2}$$

for all $s \in [t]$. If we are given an s-subset, $S \in \binom{X}{s}$, we use r_S to denote the number of blocks containing S. By standard counting techniques (e.g., see [10]), we have $r_S \geqslant \binom{v-s}{t-s}/\binom{k-s}{t-s} = r_s$. For (t,k,v) Steiner systems, this holds with equality. Since $k \leqslant v$, we have $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_{t-1} \geqslant r_t$. Additionally, the number of blocks satisfies,

$$|\mathcal{B}| \geqslant \binom{v}{t} \Big/ \binom{k}{t}, \tag{3}$$

and once again, this is attained with equality for Steiner systems. Now, define GFR codes based on covering designs:

Definition 3. *Let $(X,\mathcal{B})$ be a (t,k,v) covering design with blocks $\mathcal{B}=\{\!\{B_1,\dots,B_n\}\!\}$. A* covering fractional repetition (CFR) code *with parameters $(N=|\mathcal{B}|,k,\rho)_v$ is a $k\times N$ array $\mathcal{A}\in[v]^{k\times N}$, whose columns are the blocks of $\mathcal{B}$ in some order. Each symbol of $[v]$ appears in at least ρ columns.*

Thus, a CFR code is just the blocks of a covering design written as columns of an array. The order of the columns does not matter. Within each column we may choose to order the block elements in any way we want, however, this crucially determines the skip cost of the code.

The main relaxation of CFR codes, compared with FR codes, is that not every symbol of $[v]$ appears the same number of times in the array $\mathcal{A}$. We can always safely assume ρ from Definition 3 is at least r_1 from (2). If we denote by ρ_a the number of time $a\in[v]$ appears in $\mathcal{A}$, then we have $kN=\sum_{a\in[v]}\rho_a$, which reduces $kN=v\rho$ when all the ρ_a are equal. Thus, each node contains exactly k symbols, and each symbol $a\in[v]$ appears in ρ_a nodes. We comment that [18], [15] study the transpose design, resulting in each symbol in $[v]$ appearing the same number of times, but nodes may contain a different number of symbols.

Finally, we recall Rödl's proof of the existence of good covering designs [9]:

Lemma 1 ([9])**.** *For all positive integers $t\leqslant k\leqslant v$, there exist (t,k,v) covering designs $(X,\mathcal{B})$, with size $|\mathcal{B}|=\binom{v}{t}/\binom{k}{t}(1+o(1))$, when t and k are constants, and $v\to\infty$.*

Lemma 1 shows that it is possible to build covering designs whose number of blocks approaches that of Steiner systems, even in parameters where no Steiner systems exist.

III. Constructions of Zero Skip-Cost CFR Codes

The goal of this section is to construct CFR codes with zero skip cost. As we show shortly, this is always possible in a trivial way – simply repeat columns in the CFR code. However, we would also like to minimize the number of columns in the CFR code, a notion we shall also formalize as the *expansion factor*. Another simple construction we present gives an asymptotic optimal expansion factor, but poor expansion factor for many practical parameters. We then present a more elaborate construction which significantly reduces the expansion factor. We conclude by showing, using the probabilistic method, that in fact almost all orderings of covering designs have zero skip cost, and no expansion is needed. This, unfortunately, is not a constructive proof.

We first ask ourselves what is the locality restriction. Particularly, since lower locality is better, we shall determine the minimum locality we can guarantee given just the parameters of the covering design used to construct the CFR code. To that end we first define a set of desired design parameters.

Definition 4. *Let $(X,\mathcal{B})$ be a (t,k,v) covering design. We say it is* properly local *if*

$$r_{t-1}=\left\lceil\frac{v-t+1}{k-t+1}\right\rceil \geqslant \left\lceil\frac{k}{t-1}\right\rceil+1. \tag{4}$$

Lemma 2. *Let $\mathcal{A}\in[v]^{k\times N}$ be a CFR code based on a properly local (t,k,v) covering design $(X,\mathcal{B})$. Then the minimum locality of $\mathcal{A}$ satisfies $\ell_{\min}\leqslant\lceil k/(t-1)\rceil$.*

Following Lemma 2, we shall consider only properly local covering designs, and analyze the skip cost of CFR codes based on them using locality $\ell=\lceil k/(t-1)\rceil$. We comment that restricting ourselves to properly local covering designs is a very mild restriction, since all (v,t,k) covering designs with $v\geqslant\frac{k^2}{t-1}$ are properly local, and increasing v for any constant t and k is possible by Lemma 1.

The first construction we describe is completely trivial, serving only the purpose of providing a baseline, and motivating the definition of the expansion factor. To describe it, we define the following notation: for any multiset, S, and any integer $n\in\mathbb{N}$, we denote by $n\cdot S$ the multiset that contains n copies of each of the elements of S.

Construction 1. *Let $(X, \mathcal{B})$ be a properly local (t, k, v) covering design. We construct the CFR code, $\mathcal{A}$, whose columns are the blocks of $(X, 2\cdot\mathcal{B})$, in some arbitrary order.*

Theorem 3. *Let $\mathcal{A}$ be the CFR code from Construction 1. Then it is a $(2|\mathcal{B}|, k, r_1)_v$-CFR code with locality 1 and $\mathsf{cost}(\mathcal{A}) = 0$, where r_1 is given in (2).*

Construction 1 is indeed trivial. It also seems quite wasteful, since we simply have repeated nodes. One might therefore ask, given k, ρ, and v (or alternatively, given t, k, and v), what is the smallest N for which an $(N, k, \rho)_v$-CFR code exists. We observe that (3) gives a lower bound, which inspires the following definition:

Definition 5. *Let $\mathcal{A}$ be an $(N, k, \rho)_v$-CFR code from a (t, k, v) covering design. The* expansion factor *of $\mathcal{A}$ is defined as*

$$\xi(\mathcal{A}) \triangleq \frac{N}{\binom{v}{t}/\binom{k}{t}}.$$

We have $\xi(\mathcal{A}) \geqslant 1$ for all CFR codes, and we would obviously like to have $\xi(\mathcal{A})$ as low as possible.

Corollary 4. *For any fixed $1 \leqslant t \leqslant k$, there exists an infinite sequence of integers, $k \leqslant v_1 < v_2 < \cdots$, such that there are CFR codes, $\mathcal{A}_i$, with parameters $(N_i, k, \binom{v_i-1}{t-1}/\binom{k-1}{t-1})_{v_i}$, locality $\ell = 1$, with $\mathsf{cost}(\mathcal{A}_i) = 0$ and $\lim_{i\to\infty} \xi(\mathcal{A}_i) = 2$.*

Having constructed zero skip-cost CFR codes with asymptotic expansion factor of 2, we turn to show a construction with an asymptotic expansion factor of 1.

Construction 2. *Let $(X, \mathcal{B})$ be a properly local (t, k, v) covering design, $2 \leqslant t \leqslant k$, and let $(X, \mathcal{B}')$ be a $(t-1, k, v)$ covering design. Define $r \triangleq k \,\overline{\mathrm{mod}}\,(t-1)$.*

If $r = t-1$ or $r = 1$, we construct the CFR code, $\mathcal{A}$, whose columns are the blocks of $(X, \mathcal{B} \cup \binom{k}{t-1}\cdot\mathcal{B}')$ arranged as:

- *Blocks from $\mathcal{B}$ appear in some arbitrary order.*
- *Each block of $\mathcal{B}'$ is written in $\binom{k}{t-1}$ columns, where each possible $(t-1)$-subset of the block appears contiguously as the first $t-1$ elements in one of the $\binom{k}{t-1}$ copies.*

If $2 \leqslant r \leqslant t-2$, we construct the CFR code, $\mathcal{A}$, whose columns are the blocks of $(X, \mathcal{B} \cup \binom{k}{t-1}\cdot\mathcal{B}' \cup \binom{k}{r}\cdot\mathcal{B}')$ arranged in the following order:

- *Blocks from $\mathcal{B}$ appear in some arbitrary order.*
- *Each block of $\mathcal{B}'$ is written in $\binom{k}{t-1}$ columns, where each possible $(t-1)$-subset of the block appears contiguously as the first $t-1$ elements in one of the $\binom{k}{t-1}$ copies.*
- *Each block of $\mathcal{B}'$ is written in $\binom{k}{r}$ columns, where each possible r-subset of the block appears contiguously as the first r elements in one of the $\binom{k}{r}$ copies.*

Theorem 5. *Let $\mathcal{A}$ be the CFR code of Construction 2. Define*

$$N \triangleq \begin{cases} |\mathcal{B}| + \binom{k}{t-1}|\mathcal{B}'| & r = 1, t-1, \\ |\mathcal{B}| + \left(\binom{k}{t-1} + \binom{k}{r}\right)|\mathcal{B}'| & 2 \leqslant r \leqslant t-2. \end{cases}$$

Then $\mathcal{A}$ is a $(N, k, r_1)_v$-CFR code with locality $\ell = \lceil k/(t-1)\rceil$ and $\mathsf{cost}(\mathcal{A}) = 0$, where r_1 is given in (2).

Corollary 6. *For fixed $2 \leqslant t \leqslant k$, there is an infinite sequence of integers, $k \leqslant v_1 < v_2 < \cdots$, such that there are CFR codes, $\mathcal{A}_i$, with parameters $(N_i, k, \binom{v_i-1}{t-1}/\binom{k-1}{t-1})_{v_i}$, locality $\ell = \lceil k/(t-1)\rceil$, $\mathsf{cost}(\mathcal{A}_i) = 0$, and $\lim_{i\to\infty} \xi(\mathcal{A}_i) = 1$.*

While Construction 2 attains an expansion factor of 1, it does so only asymptotically. Thus, for many practical values its performance is inferior to the trivial Construction 1. Motivated by this, we provide a more elaborate construction of CFR codes. It is based on a method for constructing covering designs from other covering designs that is inspired by the doubling construction of Steiner quadruple systems of Hanani [4] used in [1] to get a zero skip-cost FR code. While the construction almost never results in an optimal asymptotic expansion factor, it out-performs Construction 1 and Construction 2 for many practical values.

Before proceeding, we introduce the following notation. For any vector $\boldsymbol{x} = (x_1, \ldots, x_k) \in [v]^k$ and any $a \in [v]$, we denote

$$|\boldsymbol{x}|_a \triangleq |\{i \in [k] : x_i = a\}|$$

the number of coordinates of $\boldsymbol{x}$ containing a.

Construction 3. *Let $(X, \mathcal{B})$ be a properly local (t, k, v) covering design, $X = [v]$, $\mathcal{B} \subseteq \binom{X}{k}$. Define*

$$q \triangleq \left\lceil \frac{k}{t-1} \right\rceil \qquad\qquad r \triangleq k \,\overline{\mathrm{mod}}\,(t-1). \tag{5}$$

We construct a design $(X^*, \mathcal{B}^*)$, with $X^* = [v^*]$, $\mathcal{B}^* \subseteq \binom{X^*}{k}$, where $v^* = qv$, in the following way.

We first define three index sets:

$$\begin{aligned} I_1 &\triangleq \{\boldsymbol{i} \in [q]^k : \{\!\{|\boldsymbol{i}|_1, \ldots, |\boldsymbol{i}|_q\}\!\} = \{\!\{r, \overline{t-1}\}\!\}\}, \\ I_2 &\triangleq \{\boldsymbol{i} \in [q]^k : \{\!\{|\boldsymbol{i}|_1, \ldots, |\boldsymbol{i}|_q\}\!\} = \{\!\{0, t-1+r, \overline{t-1}\}\!\}\}, \\ I_3 &\triangleq \{\boldsymbol{i} \in [q]^k : r+1 \leqslant m \leqslant \lfloor t/q \rfloor, \\ &\qquad \{\!\{|\boldsymbol{i}|_1, \ldots, |\boldsymbol{i}|_q\}\!\} = \{\!\{m, t-1+r-m, \overline{t-1}\}\!\}\}, \end{aligned}$$

where an overline means repeating the value underneath it. Given an integer $x \in X = [v]$, and an integer $i \in [q]$, we define $\langle x, i\rangle \triangleq x + (i-1)v \in X^* = [v^*]$. Similarly, for $U \subseteq X = [v]$, we define the translation of U by $(i-1)v$ as $\langle U, i\rangle \triangleq \{\langle u, i\rangle : u \in U\} \subseteq X^* = [v^*]$. We then define:

$$\begin{aligned} \mathcal{B}_1 &\triangleq \{\cup_{j=1}^{q-1}\langle U, i_j\rangle \cup \langle V, i_q\rangle : U \in \tbinom{X}{t-1}, V \in \tbinom{U}{r}, \\ &\qquad\qquad \{i_1, \ldots, i_q\} = [q]\}, \\ \mathcal{B}_2 &\triangleq \{\cup_{j=1}^{q-2}\langle U, i_j\rangle \cup \langle V, i_{q-1}\rangle \cup \langle W, i_q\rangle : \\ &\quad U \in \tbinom{X}{t-1}, V \in \tbinom{U}{m}, \\ &\quad W \in \tbinom{U}{t-1+r-m}, \{i_1, \ldots, i_q\} = [q], r+1 \leqslant m \leqslant \lfloor t/q \rfloor\}, \\ \mathcal{B}_3 &\triangleq \{\{x_1, \ldots, x_k\} \in \tbinom{X^*}{k} : \\ &\qquad (\lceil x_1/v\rceil, \ldots, \lceil x_k/v\rceil) \in I_1 \cup I_2, \\ &\qquad \{x_1 \,\overline{\mathrm{mod}}\, v, \ldots, x_k \,\overline{\mathrm{mod}}\, v\} \in \mathcal{B}\} \\ \mathcal{B}_4 &\triangleq \{\{x_1, \ldots, x_k\} \in \tbinom{X^*}{k} : (\lceil x_1/v\rceil, \ldots, \lceil x_k/v\rceil) \in I_3, \\ &\qquad \{x_1 \,\overline{\mathrm{mod}}\, v, \ldots, x_k \,\overline{\mathrm{mod}}\, v\} \in \mathcal{B}\} \end{aligned}$$

The entire block set, $\mathcal{B}^*$ is then defined by

$$\mathcal{B}^* = \begin{cases} \mathcal{B}_1 \cup \mathcal{B}_3, & \lfloor t/q \rfloor \leqslant r \leqslant t-1, \\ \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3 \cup \mathcal{B}_4, & 1 \leqslant r \leqslant \lfloor t/q \rfloor - 1. \end{cases}$$

Finally, we construct the CFR code, $\mathcal{A}$, by writing the blocks of $\mathcal{B}^*$ in columns, where the numbers in each column appear in ascending order.

Theorem 7. *Assume the setting of Construction 3. Then $(X^*, \mathcal{B}^*)$ is a (t, k, qv) covering design. Additionally,*

$$|\mathcal{B}^*| = \begin{cases} c_1(t,k)|\mathcal{B}| + O(v^{t-1}), & r = t-1, \\ c_2(t,k)|\mathcal{B}| + O(v^{t-1}), & \left\lfloor \frac{t}{q} \right\rfloor \leqslant r \leqslant t-2, \\ c_3(t,k)|\mathcal{B}| + O(v^{t-1}), & 1 \leqslant r \leqslant \left\lfloor \frac{t}{q} \right\rfloor - 1, \end{cases} \tag{6}$$

where

$$c_1(t,k) = \frac{k!}{(t-1)!^q} + q(q-1)\frac{k!}{(2t-2)!(t-1)!^{q-2}},$$
$$c_2(t,k) = q\frac{k!}{r!(t-1)!^{q-1}} + q(q-1)\frac{k!}{(t-1+r)!(t-1)!^{q-2}},$$
$$c_3(t,k) = c_2(t,k) + q(q-1)\sum_{m=r+1}^{\lfloor \frac{t}{q} \rfloor} \frac{k!}{m!(t-1+r-m)!(t-1)!^{q-2}}. \tag{7}$$

Proof sketch: The proof proceeds by induction. By slowly going over all possible cases we show every t-subset of points is contained in at least one block of $\mathcal{B}^*$. ∎

Theorem 8. *Let $\mathcal{A}$ be the CFR code from Construction 3. Then it is a $(|\mathcal{B}^*|, k, \binom{qv-1}{t-1}/\binom{k-1}{t-1})_{qv}$-CFR code with locality $\ell = \lceil k/(t-1) \rceil$ and $\mathsf{cost}(\mathcal{A}) = 0$, where $|\mathcal{B}^*|$ is given in Theorem 7.*

Corollary 9. *For any fixed $2 \leqslant t \leqslant k$, there exists an infinite sequence of integers, $k \leqslant v_1 < v_2 < \cdots$, such that there are CFR codes, $\mathcal{A}_i$, with parameters $(N_i, k, \binom{qv_i-1}{t-1}/\binom{k-1}{t-1})_{qv_i}$, locality $\ell = \lceil k/(t-1) \rceil$, with $\mathsf{cost}(\mathcal{A}_i) = 0$ and*

$$\lim_{i\to\infty} \xi(\mathcal{A}_i) = \frac{1}{q^t} \cdot \begin{cases} c_1(t,k), & r = t-1, \\ c_2(t,k), & \left\lfloor \frac{t}{q} \right\rfloor \leqslant r \leqslant t-2, \\ c_3(t,k), & 1 \leqslant r \leqslant \left\lfloor \frac{t}{q} \right\rfloor - 1, \end{cases}$$

where q and r are defined in (5)*, and $c_1(t,k)$, $c_2(t,k)$ and $c_3(t,k)$ are defined in* (7)*.*

Proof: By Lemma 1, there exists a sequence of (t, k, v_i) covering designs, $(X_i, \mathcal{B}_i)$, with

$$|\mathcal{B}| = \frac{\binom{v_i}{t}}{\binom{k}{t}}(1 + o(1)).$$

We apply Construction 3, which by Theorem 8 gives us a sequence of $(|\mathcal{B}_i^*|, k, \binom{qv_i-1}{t-1}/\binom{k-1}{t-1})$-CFR codes, $\mathcal{A}_i$, and where $|\mathcal{B}^*|$ is given by (6) and (7). We therefore have

$$\lim_{i\to\infty} \xi(\mathcal{A}_i) = \lim_{i\to\infty} \frac{|\mathcal{B}_i^*|}{\binom{qv_i}{t}/\binom{k}{t}} = \lim_{i\to\infty} \frac{|\mathcal{B}_i|}{\binom{qv_i}{t}/\binom{k}{t}} \cdot \frac{|\mathcal{B}_i^*|}{|\mathcal{B}_i|},$$

which proves our claim since the first fraction tends to $\frac{1}{q^t}$, and by (6) and (7), the second tends to $c_j(t,k)$, $j \in \{1,2,3\}$, where j is determined by the parameters ranges. ∎

TABLE II
THE ASYMPTOTIC EXPANSION FACTOR OF CFR CODES RESULTING FROM CONSTRUCTION 3 AND COROLLARY 9

(t,k)	$(3,4)$	$(4,5)$	$(4,6)$	$(5,6)$	$(5,7)$	$(5,8)$
ξ	1	$\frac{11}{8}$	$\frac{11}{8}$	1	$\frac{9}{4}$	$\frac{9}{4}$

TABLE III
THE EXPANSION FACTOR OF CFR CODES FROM $(5,6,v)$ COVERING DESIGNS USING CONSTRUCTIONS 1, 2, AND 3

v	12	14	16	18	20	22	24	26
Con. 1	2.00	2.22	2.22	2.14	2.17	2.12	2.00	2.11
Con. 2	10.32	8.30	7.37	6.03	5.52	5.03	4.32	4.21
Con. 3	1.61	1.83	1.68	1.73	1.59	1.63	1.43	1.50

Some examples of the asymptotic expansion factor resulting from Construction 3 are provided in Table II. We note that generally, the asymptotic expansion factor guaranteed by Construction 3 is worse (i.e., higher) than that of Construction 2. However, in the non-asymptotic regime, Construction 3 may significantly outperform Construction 2. As example, we consider CFR codes from $(5, 6, v)$ covering designs for even values of $12 \leqslant v \leqslant 26$. For the constructions we use the smallest known covering designs parameters given in [3]. The resulting expansion factors from Constructions 1, 2, and 3, are summarized in Table III.

We conclude this section by showing that, for all sufficiently large v, any (t, k, v) covering design may be directly placed as the columns of a CFR code with zero skip cost. This will create a CFR code with the optimal expansion factor. However, the claim is proved using the probabilistic method, and is therefore non-constructive.

Theorem 10. *Fix integers $2 \leqslant t \leqslant k$. Then for all sufficiently large v, any (t, k, v) covering design, $(X, \mathcal{B})$, can create a CFR code $\mathcal{A}$, whose columns are the blocks of $\mathcal{B}$, with locality $\ell = \lceil k/(t-1) \rceil$, such that $\mathsf{cost}(\mathcal{A}) = 0$.*

Proof sketch: We use the probabilistic method. We first place the covering design in an array. We then permute each column independently and uniformly. We bound the probability that a column does not have zero skip cost. Using the union bound we can show that with a positive probability, there exist good permutations for the columns resulting in zero skip cost. ∎

REFERENCES

[1] Y. M. Chee, S. H. Dau, T. Etzion, H. M. Kiah, Y. Luo, and W. Zhang, "Repairing with zero skip cost," in *2024 IEEE International Symposium on Information Theory (ISIT)*, 2024, pp. 2134–2139.

[2] S. El Rouayheb and K. Ramchandran, "Fractional repetition codes for repair in distributed storage systems," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2010, pp. 1510–1517.

[3] D. M. Gordon, "Covering designs," https://www.dmgordon.org/cover/, 2025.

[4] H. Hanani, “On quadruple systems,” *Canadian Journal of Mathematics*, vol. 12, pp. 145–157, 1960.

[5] Z. Liu and Z. Zhang, “A family of access-friends MDS array codes,” in *Proceedings of the 2023 IEEE International Symposium on Information Theory (ISIT2024), Athens, Greece*, Jul. 2024, pp. 2353–2358.

[6] O. Olmez and A. Ramamoorthy, “Repairable replication-based storage systems using resolvable designs,” in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2012, pp. 1174–1181.

[7] ——, “Fractional repetition codes with flexible repair from combinatorial designs,” *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 1565–1591, 2016.

[8] K. Rashmi, N. B. Shah, P. V. Kumar, and K. Ramchandran, “Explicit construction of optimal exact regenerating codes for distributed storage,” in *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2009, pp. 1243–1249.

[9] V. Rödl, “On a packing and covering problem,” *European Journal of Combinatorics*, vol. 6, no. 1, pp. 69–78, 1985.

[10] D. R. Stinson, *Combinatorial designs: constructions and analysis*. New York: Springer-Verlag, 2003.

[11] T.-Y. Wu, Y. S. Han, Z. Li, B. Bai, G. Zhang, X. Zhang, and X. Wu, “Achievable lower bound on the optimal access bandwidth of $(K+2, K, 2)$-MDS array code with degraded read friendly,” in *2021 IEEE Information Theory Workshop (ITW)*, 2021, pp. 1–5.

[12] W. Yu, B.-J. Yuan, and M. Schwartz, “On zero skip-cost generalized fractional-repetition codes from covering designs,” *arXiv preprint arXiv:2502.12897*, 2025.

[13] W. Zhang, H. M. Kiah, and S. H. Dau, “Zigzag codes revisited: from optimal rebuilding to small skip cost and small fields,” *arXiv preprint arXiv:2509.23090*, 2025.

[14] B. Zhu, “A study on universally good fractional repetition codes,” *IEEE Communications Letters*, vol. 22, no. 5, pp. 890–893, 2018.

[15] B. Zhu, H. Li, and S.-Y. R. Li, “General fractional repetition codes from combinatorial designs,” *IEEE Access*, vol. 5, pp. 26 251–26 256, 2017.

[16] B. Zhu, K. W. Shum, and H. Li, “Fractional repetition codes with optimal reconstruction degree,” *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 983–994, 2019.

[17] ——, “On the duality and file size hierarchy of fractional repetition codes,” *The Computer Journal*, vol. 62, no. 1, pp. 150–160, 2019.

[18] B. Zhu, K. W. Shum, H. Li, and H. Hou, “General fractional repetition codes for distributed storage systems,” *IEEE Communications Letters*, vol. 18, no. 4, pp. 660–663, 2014.