

Batch Codes through Dense Graphs without Short Cycles

Alexandros G. Dimakis^{*1}, Anna Gál^{§†2}, Ankit Singh Rawat^{#1}, and Zhao Song^{‡2}

¹Department of Electrical and Computer Engineering, The University of Texas at Austin, TX

[¶]dimakis@austin.utexas.edu [#]ankitsr@utexas.edu

²Department of Computer Science, The University of Texas at Austin, TX

[§]panni@cs.utexas.edu [‡]zhaos@utexas.edu

October 10, 2014

Abstract

Consider a large database of n data items that need to be stored using m servers. We study how to encode information so that a large number k of read requests can be performed *in parallel* while the rate remains constant (and ideally approaches one). This problem is equivalent to the design of multiset Batch Codes introduced by Ishai, Kushilevitz, Ostrovsky and Sahai [17].

We give families of multiset batch codes with asymptotically optimal rates of the form $1 - 1/\text{poly}(k)$ and a number of servers m scaling polynomially in the number of read requests k . An advantage of our batch code constructions over most previously known multiset batch codes is explicit and deterministic decoding algorithms and asymptotically optimal fault tolerance.

Our main technical innovation is a graph-theoretic method of designing multiset batch codes using dense bipartite graphs with no small cycles. We modify prior graph constructions of dense, high-girth graphs to obtain our batch code results. We achieve close to optimal tradeoffs between the parameters for bipartite graph based batch codes.

1 Introduction

Batch codes were introduced by Ishai, Kushilevitz, Ostrovsky and Sahai in [17], motivated by applications to load balancing in distributed storage and private information retrieval. An (n, N, k, m) batch code encodes a string \mathbf{x} of n symbols to a string \mathbf{y} of N symbols to be stored at m distinct servers (also called buckets), so that any k read requests for symbols of \mathbf{x} can be performed by reading at most 1 (more generally t) symbols from each bucket.

In this paper we study batch codes in the context of their applications in distributed storage. Consider a large database of n data items that need to be stored using m servers. The simplest and most frequently used way of introducing redundancy in distributed storage systems is replication. Assume that the whole database is replicated k times using $m = k$ servers. One advantage of k -replication is that *any* k read requests can be processed in parallel¹ This includes k requests to read one item, $k/2$ requests to read any two items, *etc.*

^{*}A. G. Dimakis would like to acknowledge support from grants NSF CCF-1422549, CCF-1344364, and CCF-1344179.

[†]A. Gál would like to acknowledge support from grant NSF CCF-1018060.

¹We assume that each server can serve only one read request at a time, but this can be easily generalized.

Further, replication provides fault tolerance since any $k - 1$ server failures can be tolerated with no data loss. The disadvantage, of course, is that the rate of this storage scheme is vanishing as $1/k$. This large storage overhead of replication is a major cost bottleneck for large-scale storage clusters that motivates the use of distributed storage codes in systems like Windows Azure and Hadoop [16, 35].

The central question is how to encode information so that a scaling number k of read requests can be performed in parallel while the rate remains constant (and ideally approaches one). This problem is equivalent to designing *multiset* batch codes with good parameters. Multiset batch codes were also introduced in [17]. These are batch codes that allow read requests that form a multiset, that is the same item x_i from \mathbf{x} may be requested multiple times. For any k (not necessarily distinct) requests x_{i_1}, \dots, x_{i_k} there must be a partition of the m servers into k disjoint groups S_1, \dots, S_k , such that the j -th request x_{i_j} can be recovered by reading at most 1 symbol from each server of the group S_j . Thus, in the context of distributed storage, multiset batch codes also allow multiple parallel reads, a very useful property when storing frequently accessed data. For example, if item x_1 is requested by k different users, it can be retrieved simultaneously, by accessing distinct groups of servers, and reading only one symbol from each server.

The central code design question for this scenario involves minimizing the number of used servers m while maximizing the rate $\rho = n/N$ for some given number of objects n and required number of parallel reads k . Clearly, for any scheme $m \geq k$. Note that $m = k$ is achievable by k -replication but with vanishing rate $\rho = 1/k$. An important contribution of [17] was the design of codes that for any constant rate $\rho < 1$ achieve polynomial dependency of the number of servers on k , that is $m = k^c$ for some constant c that depends on the rate ρ . One limitation is that this exponent c is fairly large and grows arbitrarily large as the rate ρ approaches 1. We note that for low rates [17] proposed a different construction that is near-optimal. Specifically, they showed that using Reed-Muller codes it is possible to obtain batch codes of rate $\rho = 1/k^\epsilon$ and $m = k \log^{2+1/\epsilon+o(1)} k$ for any constant $\epsilon > 0$. It remained open if it is possible to construct multiset batch codes with rates approaching optimality $\rho = 1 - o(1)$. Further, it would be ideal to achieve this using few servers, *i.e.*, $m = k^c$ for some small exponent c . By a connection between multiset batch codes and smooth codes observed by Ishai et al. (Theorem 3.9 in [17]), the recent constructions of high rate locally decodable codes by Kopparty et al. [19] imply multiset batch codes with these properties, *e.g.* rate $1 - O(\log \log k / \log k)$ achieved with $m < k^{1+O(1/\log \log k)}$, however this has not been explicitly mentioned in the literature. Other constructions of high rate locally decodable codes appear in [14, 15].

As far as we know, the only constructions of batch codes published after [17] have been replication based batch codes - referred to as *combinatorial batch codes* in the literature (see *e.g.* [1, 4–6, 26, 34, 36]). Note that combinatorial batch codes do not provide multiset batch codes.

A limitation of most previous multiset batch codes including all multiset batch codes of [17] with polynomial number of servers $m = \text{poly}(k)$ is that they do not have explicit deterministic decoders. They provide randomized decoding algorithms, and imply the existence of the disjoint groups of servers required for decoding, but do not construct the disjoint groups, and do not identify which servers to use for the parallel reads explicitly. The multiset batch codes with deterministic decoders in [17] have superpolynomial number of servers with respect to k .

Contributions: We introduce the first families of multiset batch codes with explicit and deterministic decoding algorithms and a polynomial number of servers with respect to k . Our codes also achieve asymptotically optimal rates $1 - o(1)$, and asymptotically optimal fault tolerance. Our constructions exhibit a tradeoff between m (the number of servers) and the rate as functions of k . Note that we are trying to minimize m and maximize the rate as functions of k . We obtain several constructions that give different tradeoffs. The smallest m we can achieve (as a function of k) while still having rate $1 - o(1)$ is $m = k^{3+\epsilon}$ for any fixed $\epsilon > 0$, achieved with rate at least $1 - \frac{1}{k^{\epsilon/4}}$. These tradeoffs are close to optimal for bipartite graph based

multiset batch codes. A different setting of parameters in this construction gives $m = O(k^3)$ servers and rate at least $1 - \frac{1}{c}$ for any constant $c \geq 2$. In another construction we require $m = k^{4+\epsilon}$ and achieve rate at least $1 - \frac{1}{k^\epsilon}$, for any constant $1 \geq \epsilon > 0$. The first construction gives a better tradeoff between the number of servers and the rate when ϵ is small. We also present a construction that gives $m = k^4$ achieved with rate at least $1 - \frac{1}{\sqrt{k}}$. This gives a better rate than any of the other two constructions when the number of servers $m = k^4$. We also show that the zig-zag code construction [39] allows us to design batch codes with $m = \Theta(k^{k^{1+\epsilon}+1+\epsilon})$ and rate at least $1 - \frac{1}{k^\epsilon}$ for any $\epsilon > 0$. While they require substantially more servers, zig-zag based codes offer the benefit of logarithmic locality, *i.e.*, each of the parallel reads needs to contact only $O(\text{polylog}n)$ servers in the worst case. Table 1 summarizes the parameters obtained by various code constructions considered in this paper.

Construction	Rate	Number of servers (m)
Theorem 3, Section 4.1	$\geq 1 - \frac{1}{k^{\epsilon/4}}$ (for $\epsilon > 0$)	$k^{3+\epsilon}$
Theorem 5, Section 4.2	$\geq 1 - \frac{1}{k^\epsilon}$ (for $1 \geq \epsilon > 0$)	$k^{4+\tilde{\epsilon}} + o(k^{4+\tilde{\epsilon}})$
Theorem 6, Section 4.3	$\geq 1 - 1/\sqrt{k}$	$(1 + o_k(1))k^4$
Theorem 7, Section 4.4	$\geq 1 - \frac{1}{k^\epsilon}$ (for $\epsilon > 0$)	$\Theta(k^{k^{1+\epsilon}+1+\epsilon})$
Theorem 4, Section 4.1	$\geq 1 - \frac{1}{c}$ (for constant $c \geq 2$)	$O(k^3)$

Table 1: Summary of the constructions of multiset batch codes presented in this paper. For the multiset batch codes constructed in Section 4.2, $\tilde{\epsilon} < \epsilon$ is a rational number which depends on ϵ . Note that we have deterministic decoding algorithm for all our constructions.

Our constructions offer several benefits compared to prior work. The first is in decoding: the choice of which servers to use for each of the k parallel reads is deterministic and obtained using an efficient algorithm. This is in contrast with previous constructions of multiset batch codes with polynomial number of servers with respect to k that choose these sets randomly and only show the existence of assignments of disjoint groups of servers to read requests. The second benefit is that our multiset batch codes achieve rate of the form $1 - \frac{1}{\text{poly}(k)}$ with a polynomial number of servers, which is larger than the rate $1 - O(\log \log k / \log k)$ of batch codes implied by the recent high rate locally decodable codes of [19]. Finally, our batch codes offer fault tolerance. Specifically, we show how to augment the code with extra parity symbols to provide distance asymptotically approaching that of Maximum Distance Separable (MDS) codes. We note that while this transformation can be applied to any batch code, batch codes with rate $1 - o(1)$ are necessary to obtain asymptotically optimal fault tolerance.

Techniques: Our main technical innovation is a graph-theoretic method of designing multiset batch codes using dense bipartite graphs with no short cycles. Consider first the case where each server stores one symbol (*i.e.* $N = m$), called a *primitive* batch code. Primitive batch codes are subsequently used as building blocks to construct more general batch codes.

Consider a bipartite graph with n left nodes and $N - n$ right nodes. Each left node corresponds to a data symbol and each right node to a parity symbol. Let the left nodes all have degree k , indicating that each data symbol x_i is included in k parities. It is not hard to see that if this bipartite graph has no 4-cycle, these k parities must have disjoint neighborhoods, excluding x_i . Therefore, if there are no 4-cycles, each symbol x_i can be reconstructed k times in parallel, by reading these parities and their disjoint neighborhoods.

This simple observation allows k parallel reads for one symbol but cannot support for example $k/2$ reads for x_1 and $k/2$ reads for x_2 since all the parities of x_1 and x_2 might involve other common data symbols. Our main lemma shows that if there are no 4-cycles or 6-cycles, the number of parities that are blocked by each read request can be bounded by a local argument and any k reads can be supported. Using this graph theoretic lemma we show that bipartite graphs with n left nodes, $N - n$ right nodes, left degree at least k and girth at least 8 yield primitive multiset batch codes with parameters (n, N, k, m) . Therefore optimizing batch code parameters becomes an extremal graph problem: Maximize the left minimum degree in a (simple, undirected) bipartite graph with N vertices so that it contains no 4-cycle or 6-cycle subgraphs. Here subgraphs are not necessarily induced and this condition is equivalent to the bipartite graph having girth (length of shortest cycle) at least 8.

The difficulty in code constructions is now clear: making the left degree k large without short cycles requires many right nodes. However, as the number of right nodes (*i.e.* parities) is $N - n$, this pushes the rate n/N low. The question of constructing dense bipartite graphs of high girth and in particular bipartite graphs without 4- and 6-cycles and as many edges as possible has been extensively studied. The strongest constructions of such graphs have been given by Lazebnik, Ustimenko, Woldar [20] and by de Caen and Székely [7]. The construction by Lazebnik *et al.* is obtained by embedding Chevalley group geometries in Lie algebras [41]. The constructions of de Caen and Székely are based on known constructions of generalized quadrangles [27, 40]. We modify a construction of Balbuena [3] based on Latin Squares to obtain our results with $m = k^{3+\epsilon}$ and rate at least $1 - \frac{1}{k^{\epsilon/4}}$. This gives close to optimal tradeoffs for bipartite graph based multiset batch codes, since as we discuss in Section 4.5 rate at least $1 - \frac{1}{k^{\epsilon/4}}$ implies $m = \Omega(k^{3+\epsilon/2})$. Note that the graphs of Balbuena [3] only give rate $1/2$, but we apply a simple transformation to amplify the rate. We could apply a similar transformation to any of our bipartite graph based constructions, but these would not lead to significantly different tradeoffs.

1.1 Related work

Our work is inspired by work of Ishai *et al.* [17]. Batch codes were originally motivated for load balancing applications which are traditionally different from erasure codes like Reed-Solomon [32] and information dispersal [29] used for distributed data storage [28]. The difference is that batch codes are useful for storing many small items while standard erasure codes provide fault tolerance for a single large object split in multiple symbols.

One issue that blurs this difference, however, is the so-called repair problem of distributed storage codes [10, 11]. When a systematic code is used for storage across servers a very common operation is the access of a single systematic symbol. Typically, each symbol of a code is stored in a different server. When the particular server storing the desired systematic symbol has failed or is temporarily unavailable the problem of reconstructing a single symbol from as few other codeword symbols as possible arises. The locality of a symbol [13, 25] is the smallest number of other symbols that need to be read to reconstruct it. We refer to that set of symbols as a repair group for the given symbol. Batch codes enable multiple (k) repair groups for each symbol that are also disjoint and hence allow these reads to be performed in parallel, a property very useful for high performance storage systems. Since these groups must be disjoint, some of them must be small so batch codes implicitly enforce locality. The weaker problem of enabling multiple parallel reads of a single symbol was recently investigated in [31]. The constructions we present in this paper resolve an open question stated in [31].

Fault tolerance is another useful property for distributed storage codes. Gopalan *et al.* [13] determined the largest possible distance for codes with locality, generalizing the Singleton bound [23]. This was re-

cently further generalized for nonlinear codes [25] and several explicit constructions of locally repairable codes have been investigated (*e.g.* [18, 24, 30, 37, 38]). Note that locally repairable codes are related but fundamentally different from locally decodable codes [42]. Prior batch code constructions with good rates have very poor distance properties, namely $d = o(n)$. We construct batch codes that have near-optimal distance, linear in n and asymptotically approaching the Singleton bound.

We note that girth and graph representations have been considered before (see *e.g.* [1, 26] in the context of 2-uniform combinatorial batch codes, that is replication based batch codes where each message symbol is repeated exactly twice. However, this connection is very different from our technical contributions. 2-uniform combinatorial batch codes can be viewed as graphs the following way: since each symbol is stored at exactly two servers, one can represent these codes by a graph whose vertex set is the set of servers, and an edge is placed for each symbol, connecting the two servers where the symbol is stored. Note that combinatorial batch codes repeating symbols more than twice give rise to hypergraphs. In the case of 2-uniform combinatorial batch codes, the corresponding graph cannot contain certain configurations as subgraphs, and in particular the girth of the graph is also relevant. [1] uses a construction of Lazebnik, Ustimenko, Woldar [21] of high girth graphs to construct 2-uniform combinatorial batch codes. However, this is a different construction and a different paper than the construction in [20] that we use, and our graph representation and connection between high girth graphs and multiset batch codes is completely different.

1.2 Organization of the paper

In Section 2, we provide necessary background on batch codes, including the gadget lemma [17]. The gadget lemma allows us to convert a primitive batch code to a general batch code without the restriction that the number of servers be equal to the number of code symbols. We describe our main technical contribution connecting dense high-girth graphs to multiset batch codes in Section 3. We first associate a natural bipartite graph representation \mathcal{G}_C with a systematic linear code C in Section 3.1, and define the notion of repair groups for such codes in Section 3.2. In Section 3.3, we establish that having an induced subgraph of \mathcal{G}_C with left degree at least k and girth at least 8 is a sufficient condition for the code to be a primitive multiset batch code, which can support any sequence of k reads requests. In Section 4 we use this result with various explicit constructions of dense bipartite graphs with girth at least 8 to obtain multiset batch codes with rate $1 - o(1)$. In Section 4.5 we discuss the limitations of our bipartite graph based approach. In Section 5, we combine the load balancing aspect of batch codes with fault tolerance, and describe a simple way to obtain batch codes that are arbitrarily close to the Singleton bound.

Notation: A quick note on the notation that we follow throughout the paper. We use boldface lower case letters to denote vectors. For a strictly positive integer n , $[n]$ denotes the set $\{1, 2, \dots, n\}$. For two integers $0 \leq n_1 \leq n_2$, we use $[n_1 : n_2]$ to represent the set $\{n_1, n_1 + 1, \dots, n_2\}$. The notations which are specific to various sections of the paper are defined where we use them.

2 Background

Let Σ denote a finite alphabet² and $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \Sigma^n$ represent a length- n message vector. We use the (multi-)set $\{i_1, i_2, \dots, i_k\}$ to represent the sequence of k read requests that correspond to the message symbols $x_{i_1}, x_{i_2}, \dots, x_{i_k}$. Note that a multiset with k elements from $[n]$ denotes a sequence of k read requests where some message symbols may be requested multiple times. In what follows, we also refer to a sequence of k read requests as a *k-request pattern*. We start with the definition of batch codes [17]:

²We subsequently assume Σ to be a finite field.

Definition 1. (Batch codes [17]) An (n, N, k, m, t) batch code encodes n message symbols over Σ to N code symbols over Σ . These N code symbols are partitioned into m buckets (servers) such that any k -request pattern represented by a set with k distinct elements from $[n]$, i.e., parallel reads for k distinct message symbols, can be served by downloading at most t code symbols from each of the m buckets.

The notion of batch codes is further strengthened into *multiset batch codes* which also support the k -request patterns represented by multisets with k elements from the set $[n]$:

Definition 2. (Multiset batch codes [17]) An (n, N, k, m, t) multiset batch code encodes n message symbols over Σ to N code symbols over Σ . These N code symbols are partitioned into m buckets such that any k -request pattern represented by a multiset with k elements from $[n]$ can be served by downloading at most t code symbols from each of the m buckets. In addition, all read requests are required to be served from the code symbols downloaded from disjoint sets of buckets.

Remark 1. A particular sub-class of batch codes where the number of buckets is equal to the length of codewords, i.e., $m = N$, is referred to as *primitive batch codes*. Note that primitive batch codes correspond to batch codes with each of their buckets comprising exactly one code symbol.

Remark 2. In this paper we focus on batch codes with $t = 1$, i.e., at most only one symbol can be read from each of the buckets during the process of serving a k -request pattern. Given an $(n, N, k, m, 1)$ batch code, one can obtain an (n, N, tk, m, t) batch code and an $(n, N, k, \lceil \frac{m}{t} \rceil, t)$ batch code [17, Lemma 2.4.2]. In what follows, we refer to an $(n, N, k, m, 1)$ batch code as an (n, N, k, m) batch code.

In [17], Ishai *et al.* highlight a basic operation to convert a primitive batch code to a general batch code with suitable parameters. The following result describes this operation.

Lemma 1. (Gadget lemma [17]) Given an (n, N, k, m) multiset batch code \mathcal{C} , for any $g \in \mathbb{N}$, it is possible to construct a (gn, gN, k, m) multiset batch code $g \cdot \mathcal{C}$.

For completeness, we include a proof of the gadget lemma.

Proof. We prove the Lemma by presenting an explicit construction for a (gn, gN, k, m) multiset batch code $g \cdot \mathcal{C}$ which utilizes the given (n, N, k, m) multiset batch code \mathcal{C} . Let $\mathbf{x} = (x_1, x_2, \dots, x_{gn})$ be a message vector of length gn . We partition the message vector into g sub-vectors $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^g$ such that

$$\mathbf{x}^i = (x_{(i-1)n+1}, x_{(i-1)n+2}, \dots, x_{in}), \text{ for } i \in [g].$$

Assuming that $\mathcal{C}(\mathbf{x}^i)$ denotes the length- n codeword corresponding to \mathbf{x}^i in \mathcal{C} , we assign the following codeword in $g \cdot \mathcal{C}$ to the message vector \mathbf{x} .

$$g \cdot \mathcal{C}(\mathbf{x}) = \mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_N).$$

Here, for $j \in [N] = [m]$, $\mathbf{c}_j = (\mathcal{C}(\mathbf{x}^1)_j, \mathcal{C}(\mathbf{x}^2)_j, \dots, \mathcal{C}(\mathbf{x}^g)_j)$ represents the j -th bucket associated with the codeword $\mathbf{c} \in g \cdot \mathcal{C}(\mathbf{x})$. We now establish that $g \cdot \mathcal{C}$ is indeed a (gn, gN, k, m) multiset batch code.

It is clear from the construction described above that the codewords in $g \cdot \mathcal{C}$ are obtained by stacking up g codewords in \mathcal{C} , one for each of the g sub-vectors $\{\mathbf{x}^i\}_{i \in [g]}$. Given this structure of the codewords in $g \cdot \mathcal{C}$, we can associate any k -request pattern for \mathbf{x} to a k -request pattern for $\mathbf{x}^1 = (x_1, x_2, \dots, x_n)$: for any $j \in [n]$, we map a read request for the message symbol $x_{(i-1)n+j}$ to a read request for the message symbol x_j . For $j \in [n]$, let k_j denote the number of read requests from the k -request pattern for \mathbf{x} that are mapped to the read requests for the message symbol x_j . Note that we have $\sum_{j=1}^n k_j = k$. Since \mathcal{C} is an (n, N, k, m) primitive batch code, we can assign the k read requests for the symbols in the sub-vector \mathbf{x}^1 to disjoint sets of buckets. Now, for any $j \in [n]$, one can serve all k_j read requests mapped to x_j by reading the appropriate code symbols from the assigned sets of buckets. \square

3 Obtaining Multiset Batch Codes from Bipartite Graphs

This section presents the primary technical contribution of this paper. Here, we focus on systematic linear codes and provide a sufficient condition for such a code to be a primitive multiset batch code. For the remainder of the paper, we assume Σ to be a finite field over which our linear codes are defined. In particular, the parity symbols of our codes correspond to linear combinations of the message symbols with coefficients chosen from Σ .

For every systematic linear code, we associate a bipartite graph with the code. We then show that an $[n, N]$ systematic code³ is an $(n, N, k, m = N)$ primitive multiset batch code if the associated bipartite graph has an induced subgraph with girth at least 8 and left degree at least k . The proof of the fact that a systematic code which satisfies the aforementioned graph theoretic conditions is a multiset batch code has an added advantage. The proof involves showing that for such codes any k -request pattern can be assigned to disjoint sets of code symbols by a greedy and efficient algorithm (see Figure 4).

3.1 Graph representation for systematic codes

Given an $[n, N]$ systematic code \mathcal{C} , we define a bipartite graph $\mathcal{G}_{\mathcal{C}} = (\mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{E})$ with $|\mathcal{V}_1| = n$ and $|\mathcal{V}_2| = N - n$. The n left nodes and the $N - n$ right nodes in the graph $\mathcal{G}_{\mathcal{C}}$ represent the n message symbols and the $N - n$ parity symbols of the code \mathcal{C} , respectively. Moreover, there exists an edge $(i, j) \in \mathcal{E} \subseteq \mathcal{V}_1 \times \mathcal{V}_2$ iff the i -th message symbol participates in the j -th parity symbol. Alternatively, given $G = [I_{n \times n} \mid E_{n \times (N-n)}]$ as the $n \times N$ generator matrix of the code \mathcal{C} , the matrix E is supported on the incidence matrix of the bipartite graph $\mathcal{G}_{\mathcal{C}}$.

Example 1. Let \mathcal{C} be a $[3, 8]$ systematic code which encodes the message vector $\mathbf{x} = (x_1, x_2, x_3)$ to a length-8 codeword \mathbf{c} such that

$$\mathbf{c} = (c_1, c_2, \dots, c_8) = (x_1, x_2, x_3, x_1 + x_2, x_1, x_2 + x_3, x_1 + x_3, x_1 + x_2 + x_3).$$

Figure 1 describes the graph representation for \mathcal{C} .

Remark 3. Note that the graph representation considered here is different from the prevalent Tanner graph representation of the parity check matrix of a code [33]. In the graph $\mathcal{G}_{\mathcal{C}}$, all nodes represent the symbols of a generic codeword with the left nodes and the right nodes representing the message and the parity symbols, respectively. On the other hand, the *variable* and *check* nodes in a Tanner graph represent the code symbols and the different constraints satisfied by the codewords, respectively. However, there is a correspondence between these two graph representations for a systematic code as one can be obtained from the other.

3.2 Repair groups of a systematic code

We now define the notion of repair groups for a systematic linear code \mathcal{C} . For $i \in [n]$, let $c_{i_1}, c_{i_2}, \dots, c_{i_{u(i)}}$ denote the $u(i)$ parity symbols that involve the message symbol x_i . For $j \in \{i_1, i_2, \dots, i_{u(i)}\}$, we use $\mathbf{x}_{\Gamma(i,j)} \subseteq \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$ to represent the collection of the message symbols other than x_i that participate in the construction of the parity symbol c_j . This implies that for each $j \in \{i_1, i_2, \dots, i_{u(i)}\}$, one

³An $[n, N]$ systematic code is a linear code over finite field Σ that encodes n message symbols over Σ to length- N codewords over Σ . Moreover, without loss of generality we assume that the n message symbols (systematic part) appear at the first n positions of the codewords in \mathcal{C} .

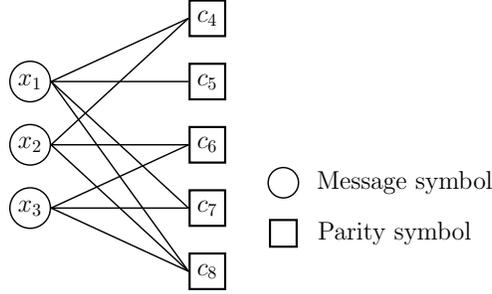


Figure 1: Graph representation \mathcal{G}_C for the $[3, 8]$ code described in Example 1.

can obtain x_i by removing the contributions of the symbols in $\mathbf{x}_{\Gamma(i,j)}$ from c_j . Noting this, we refer the $u(i)$ collections of the symbols $\{\mathbf{x}_{\Gamma(i,j)} \cup \{c_j\}\}_{j \in \{i_1, i_2, \dots, i_{u(i)}\}}$ as $u(i)$ repair groups for the message symbol x_i .

In terms of the graph representation \mathcal{G}_C , for $i \in [n]$, the left node x_i has degree $u(i)$. The right nodes associated with the parity symbols $c_{i_1}, c_{i_2}, \dots, c_{i_{u(i)}}$ constitute the $u(i)$ neighbors of the left node x_i . Moreover, for $j \in \{i_1, i_2, \dots, i_{u(i)}\}$, the left nodes associated with the message symbols $\mathbf{x}_{\Gamma(i,j)}$ correspond to all the neighbors of the right node c_j other than the left node x_i .

For the $[3, 8]$ code described in Example 1 (see Figure 1), the message symbol x_1 participates in four parity symbols c_4, c_5, c_7, c_8 . Therefore, the message symbol x_1 has the following four repair groups,

$$\{x_2, c_4\}, \{c_5\}, \{x_3, c_7\}, \{x_2, x_3, c_8\}.$$

The repair groups for the message symbols x_2 and x_3 can be listed in a similar manner.

3.3 Sufficient condition for multiset batch codes

As our main technical contribution, the following result presents a sufficient condition for a linear systematic code to be a primitive multiset batch code.

Theorem 1. *Let \mathcal{C} be an $[n, N]$ systematic code with the graph representation $\mathcal{G}_C = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{E})$. Assume that there exists an induced subgraph $\tilde{\mathcal{G}}_C = (\mathcal{V}_1, \tilde{\mathcal{V}}_2 \subseteq \mathcal{V}_2, \tilde{\mathcal{E}})^4 \subseteq \mathcal{G}_C$ such that the following two conditions hold.*

- (i) *Each node in \mathcal{V}_1 has degree at least k in the bipartite graph $\tilde{\mathcal{G}}_C$.*
- (ii) *The bipartite graph $\tilde{\mathcal{G}}_C \subseteq \mathcal{G}_C$ has girth (length of the shortest cycle) at least 8.*

Then, \mathcal{C} is an $(n, N, k, m = N)$ primitive multiset batch code.

Before proceeding to establish Theorem 1, we present Lemma 2 and 3. These two lemmas characterize the interaction among those repair groups of the message symbols that are associated with a suitable induced subgraph $\tilde{\mathcal{G}}_C \subseteq \mathcal{G}_C$. Note the requirement (i) in Theorem 1 requires the graph to be dense, while the requirement (ii) demands that the graph be not too dense.

Lemma 2. *Let \mathcal{C} be an $[n, N]$ systematic code with the graph representation $\mathcal{G}_C = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{E})$. Assume that there exists an induced subgraph $\tilde{\mathcal{G}}_C = (\mathcal{V}_1, \tilde{\mathcal{V}}_2 \subseteq \mathcal{V}_2, \tilde{\mathcal{E}}) \subseteq \mathcal{G}_C$ such that the following two conditions hold.*

⁴For $\tilde{\mathcal{V}}_2 \subseteq \mathcal{V}_2$, the edge set associated with $\tilde{\mathcal{G}}_C$ is defined as $\tilde{\mathcal{E}} := \{(i, j) \in \mathcal{E} : j \in \tilde{\mathcal{V}}_2\}$.

- (i) Each node in \mathcal{V}_1 has degree at least k in the bipartite graph $\tilde{\mathcal{G}}_C$,
- (ii) The bipartite graph $\tilde{\mathcal{G}}_C \subseteq \mathcal{G}_C$ has no 4-cycle, i.e., cycle of length 4.

Then, each message symbol has at least k disjoint repair groups.

Proof. Let $\tilde{\mathcal{C}}$ denote the systematic code associated with the subgraph $\tilde{\mathcal{G}}_C \subseteq \mathcal{G}_C$. Without loss of generality, we assume that $\tilde{\mathcal{V}}_2 \subseteq \mathcal{V}_2$ corresponds to the first $|\tilde{\mathcal{V}}_2|$ out of $|\mathcal{V}_2| = N - n$ right nodes of the graph \mathcal{G}_C ⁵. Then, every codeword $\mathbf{c} = (c_1, c_2, \dots, c_N) \in \mathcal{C}$ is associated to a unique codeword $\tilde{\mathbf{c}} = (c_1, c_2, \dots, c_{n+|\tilde{\mathcal{V}}_2|}) \in \tilde{\mathcal{C}}$ obtained by removing the last $|\mathcal{V}_2| - |\tilde{\mathcal{V}}_2|$ code symbols from the codeword \mathbf{c} . In the rest of the proof we focus only on the repair groups that contain the parity symbols associated with the set $\tilde{\mathcal{V}}_2$. In other words, we work with the code $\tilde{\mathcal{C}}$.

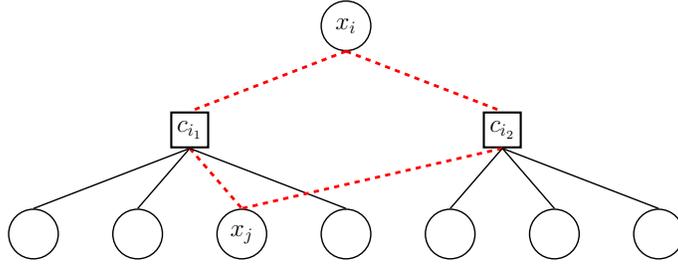


Figure 2: Illustration of a 4-cycle $\{x_i, c_{i_1}, x_j, c_{i_2}\}$ when two repair groups of the message symbol x_i have a common message symbol x_j .

Since each left node has degree at least k in the bipartite graph $\tilde{\mathcal{G}}_C$, this implies that for any $i \in [n]$, the message symbol x_i has at least k repair groups (see Section 3.2). Next, we argue that all of these repair groups are disjoint, i.e., they do not have any common code symbol. In order to establish contradiction, we assume that for an $i \in [n]$, there exists two repair groups of the message symbol x_i with at least one common code symbol. Since two repair groups for a message symbol involve distinct parity symbols, let c_{i_1} and c_{i_2} denote the parity symbols in these two intersecting repair groups. Assume that x_j , for $j \in [n] \setminus \{i\}$, be one of the common message symbol in the two underlying repair groups. This implies that the nodes $\{x_i, c_{i_1}, x_j, c_{i_2}\}$ form a cycle in the graph $\tilde{\mathcal{G}}_C$ (see Figure 2). This leads to contradiction as the graph $\tilde{\mathcal{G}}_C$ has no 4-cycle. \square

Lemma 3. Let \mathcal{C} be an $[n, N]$ systematic code with the graph representation $\mathcal{G}_C = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{E})$. Assume that there exists an induced subgraph $\tilde{\mathcal{G}}_C = (\mathcal{V}_1, \tilde{\mathcal{V}}_2 \subseteq \mathcal{V}_2, \tilde{\mathcal{E}}) \subseteq \mathcal{G}_C$ such that the following two conditions hold.

- (i) Each node in \mathcal{V}_1 has degree at least k in the bipartite graph $\tilde{\mathcal{G}}_C$,
- (ii) The bipartite graph $\tilde{\mathcal{G}}_C \subseteq \mathcal{G}_C$ has girth (length of the shortest cycle) at least 8.

Then, each message symbol has at least k disjoint repair groups. Moreover, for any $i, j \in [n]$ with $i \neq j$, any one of the disjoint repair groups for the message symbol x_i has common symbols with at most one of the disjoint repair groups for the message symbol x_j .

Proof. Here, we use ideas which are similar to those employed in the proof of Lemma 2. In particular, we associate a code $\tilde{\mathcal{C}}$ to the subgraph $\tilde{\mathcal{G}}_C \subseteq \mathcal{G}_C$. Note that for every codeword $\mathbf{c} = (c_1, c_2, \dots, c_N) \in \mathcal{C}$,

⁵If this is not the case, we can re-index the parity symbols of the codewords in \mathcal{C} to get this property.

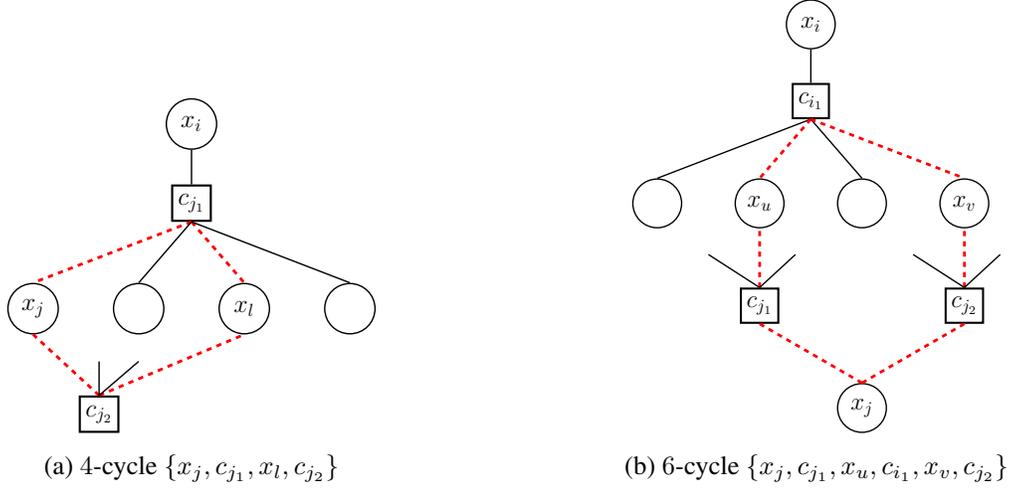


Figure 3: Illustration of the cycles in $\tilde{\mathcal{G}}_{\mathcal{C}}$ when one repair group of the message symbol x_i intersects with two disjoint repair groups of the message symbol x_j .

there exists a unique codeword $\tilde{\mathbf{c}} = (c_1, c_2, \dots, c_{n+|\tilde{\mathcal{V}}_2|}) \in \tilde{\mathcal{C}}$. Again in the rest of the proof we only focus on those repair groups that are associated with $\tilde{\mathcal{C}}$, *i.e.*, the repair groups that do not contain the codes symbols $c_{n+|\tilde{\mathcal{V}}_2|+1}, \dots, c_n$ from a codeword $\mathbf{c} = (c_1, \dots, c_N) \in \mathcal{C}$.

Since the sub graph $\tilde{\mathcal{G}}_{\mathcal{C}}$ has girth at least 8, it does not contain a 4-cycle. It then follows from Lemma 2 that each of the n message symbols has at least k disjoint repair groups. Next, we establish the second claim in the lemma that for any $i, j \in [n]$ with $i \neq j$, any one of the disjoint repair groups for the message symbol x_i has common symbols with at most one of the disjoint repair groups for the message symbol x_j .

Toward this, we assume that the opposite is true and there exist two message symbols, say x_i and x_j , such that at least one repair group of the message symbol x_i has common code symbols with at least two repair groups of the message symbol x_j . Let c_{i_1} be the parity symbol corresponding to one such repair groups of x_i that intersects with two repair groups of the message symbol x_j . Assume that c_{j_1} and c_{j_2} represent the distinct parity symbols associated with the underlying two repair groups the message symbol x_j . There are two cases that need to be considered:

- $i_1 = j_1$ **or** $i_1 = j_2$: Without loss of generality, we assume that $i_1 = j_1$. Let x_l be the message symbol that is common between the repair group of x_i and the repair group of x_j corresponding to the parity symbol c_{j_2} . As described in Figure 3a, this implies the existence of a 4-cycle $\{x_j, c_{j_1}, x_l, c_{j_2}\}$ in the bipartite graph $\tilde{\mathcal{G}}_{\mathcal{C}}$.
- $i_1 \neq j_1$ **and** $i_1 \neq j_2$: In this case, let x_u represent a common symbol between the repair group of x_i and the *first repair group* of x_j corresponding to the parity symbol c_{j_1} . Similarly, assume that x_v denotes a symbol which is common between the repair group of x_i and the *second repair group* of x_j corresponding to the parity symbol c_{j_2} . This translates to a 6-cycle $\{x_j, c_{j_1}, x_u, c_{i_1}, x_v, c_{j_2}\}$ in the bipartite graph $\tilde{\mathcal{G}}_{\mathcal{C}}$ (see Figure 3b).

Since the girth of the bipartite graph $\tilde{\mathcal{G}}_{\mathcal{C}}$ is at least 8, both cases above lead to contradiction. This completes the proof. □

Essentially, Lemma 3 implies that given an induced subgraph $\tilde{\mathcal{G}}_{\mathcal{C}}$ with girth at least 8, among the repair groups associated with $\tilde{\mathcal{G}}_{\mathcal{C}}$, each repair group of one message symbol can block at most one repair group of any other message symbol. In the following, we utilize this to prove Theorem 1.

Proof of Theorem 1. Given the code \mathcal{C} , we assign each symbol from the codeword $\mathbf{c} \in \mathcal{C}$ to a different bucket. This implies that $m = N$. Here, we provide an algorithmic proof of the fact that \mathcal{C} can support any k -request pattern by serving different read requests from the disjoint sets of buckets. In particular, under the assumptions on the bipartite graph $\tilde{\mathcal{G}}_{\mathcal{C}}$ stated in Theorem 1, we provide an algorithm which always maps k read requests in any k -request pattern to k disjoint sets of code symbols (buckets). The algorithm described here is deterministic as opposed to the randomized algorithm used for Subset codes by Ishai *et al.* in [17].

Let $\{i_1, i_2, \dots, i_k\} \subset [n]$ be a multiset corresponding to a k -request pattern. For the k -request pattern and $j \in [n]$, let k_j denote the number of read requests for the message symbol x_j in the k -request pattern, *i.e.*,

$$k_j = |\{u \in [k] : i_u = j\}|, \text{ for } j \in [n].$$

Note that $\sum_{j=1}^n k_j = k$. Moreover, each k -request pattern can be alternatively defined by a set

$$\{(1, k_1), (2, k_2), \dots, (n, k_n)\} \in [n] \times [k],$$

where (j, k_j) implies that there are k_j read requests for the message symbol x_j in the k -request pattern. We now use the algorithm described in Figure 4 to map the k read requests to disjoint sets of code symbols (buckets).

It follows from Lemma 3 that each message symbol has at least k disjoint repair groups in \mathcal{C} . Let $j^* = \min\{j \in [n] : k_j > 0\}$. The algorithm can successfully assign the k_{j^*} read requests for the message symbol x_{j^*} to the buckets storing x_{j^*} itself and $k_{j^*} - 1$ disjoint repair groups of x_{j^*} . In order to complete the proof of Theorem 1, it is enough to show that for any $j > j^*$, the algorithm is always able to find k_j sets of disjoint code symbols (buckets) to serve k_j read requests for the message symbol x_j .

Towards this, we assume that for a $j > j^*$, $k_j > 0$ and the algorithm is able to map read requests

$$\{(1, k_1), (2, k_2), \dots, (j-1, k_{j-1})\}$$

to disjoint sets of code symbols. Let $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_{\sum_{u=1}^{j-1} k_u}$ denote the disjoint sets of code symbols used by the algorithm to serve these $\sum_{u=1}^{j-1} k_u$ read requests. Note that the set \mathcal{S}_v , for $v \in [\sum_{u=1}^{j-1} k_u]$, corresponds to a message symbol in $\{x_1, x_2, \dots, x_{j-1}\}$ or a repair group of a message symbol in $\{x_1, x_2, \dots, x_{j-1}\}$. It follows from Lemma 3 that the message symbol x_j has at least k disjoint repair groups. Moreover, each of the sets $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_{\sum_{u=1}^{j-1} k_u}$ has common code symbols with at most one repair group for the message symbol x_j . Therefore, the algorithm can find $k_j \leq (k - \sum_{u=1}^{j-1} k_u)$ repair groups for the message symbol x_j which consist of code symbols that do not appear in the sets $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_{\sum_{u=1}^{j-1} k_u}$. \square

4 New Constructions of Multiset Batch Codes

We now utilize Theorem 1 to obtain new constructions of (n, N, k, m) multiset batch codes with rate $1 - o_k(1)$ ⁶. Note that the requirement (i) in Theorem 1 translates to the graph being dense, while the requirement (ii) demands that the graph be not too dense. Here, we also point out that for \mathcal{C} to be an $(n, N, k, m = N)$ primitive multiset batch code, it is necessary for each node in \mathcal{V}_1 to have degree at least k in the original

⁶Here, $b = o_k(a)$ implies that $\frac{b}{a} \rightarrow 0$ as $k \rightarrow \infty$

Algorithm: Assignment of k read requests to k disjoint sets of code symbols (buckets).

Input: An $[n, N]$ systematic code \mathcal{C} that satisfies the assumptions stated in Theorem 1, and k -request pattern $\{(1, k_1), (2, k_2), \dots, (n, k_n)\}$.

```

1:  $j = 1, \lambda = 0$ .
2: while  $\sum_{l=j}^k k_l > 0$  do
3:   if  $k_j > 0$  and the code symbol  $x_j$  is not used to serve read requests for  $x_1, x_2, \dots, x_{j-1}$  then
4:     Serve one of the  $k_j$  read requests for  $x_j$  using the bucket storing the message symbol  $x_j$ .
5:      $\lambda = 1$ . /*  $\lambda = 1$  indicates that 1 request for  $x_j$  has been served. */
6:   end if
7:   Select  $k_j - \lambda$  disjoint repair groups for  $x_j$  that do not contain any code symbol used to serve the read requests for  $x_1, x_2, \dots, x_{j-1}$ .
8:   Map  $k_j - \lambda$  (remaining) read requests for  $x_j$  to the selected  $k_j - \lambda$  disjoint repair groups.
9:    $j = j + 1$ .
10:   $\lambda = 0$ .
11: end while

```

Figure 4: Description of the algorithm to serve k read requests by using disjoint sets of code symbols (buckets).

graph $\mathcal{G}_{\mathcal{C}}$. In this section we take known explicit constructions of *dense* bipartite graphs with *girth at least 8* and study the parameters of multiset batch codes obtained from these graphs. These graphs strike a balance between the two counteracting requirements in Theorem 1.

4.1 Constructions based on the Balbuena graphs [3]

Balbuena [3] proposed a method based on Latin Squares to construct the adjacency matrices of regular bipartite graphs of girth 8. The construction of these graphs involves the concept of quasi row-disjoint matrices from [2]. The main result in [3] can be summarized as follows:

Proposition 1. (*[3]*) *Let q be a prime power and w be an integer such that $3 \leq w \leq q$. Then, there exists an explicit w -regular balanced bipartite graph of girth 8 on $2(wq^2 - q)$ nodes.*

The bipartite graphs from [3] along with Theorem 1 yield multiset batch codes with rate $1/2$ and the following parameters.

Theorem 2. *For any prime power k , and a large enough integer n , there is an explicit (n, N, k, m) Batch code with rate $\frac{1}{2}$, $m = 2k^3 - 2k$, and*

$$N = \left\lceil \frac{n}{k^3 - k} \right\rceil m$$

Proof. Apply Proposition 1 and set $w = q = k$ to obtain an $(n_1 = k^3 - k, N_1 = 2(k^3 - k), k, m_1 = N_1)$ primitive multiset batch code. Now, we can complete the proof by applying the gadget lemma. \square

As we have seen, using the graphs from [3] directly, we obtain multiset batch codes with rate $1/2$. We can modify the graph construction of [3], and obtain multiset batch codes with rate $1 - o_k(1)$ as follows.

Theorem 3. For any $\epsilon > 0$, any large enough prime power q , and large enough integer n , there is an explicit (n, N, k, m) multiset batch code with rate at least $1 - \frac{1}{k^{\epsilon/4}}$, $m = k^{3+\epsilon} + o(k^{3+\epsilon})$, and

$$N = \left\lceil \frac{n}{k^{3+\epsilon}} \right\rceil m$$

where $k = \lfloor \frac{q}{b} \rfloor$, $b = \lceil (q^3 - q)^\beta \rceil$ and $\beta = \frac{1}{3} \frac{\epsilon}{4+\epsilon}$.

Proof. We start with the 4 and 6-cycle free regular bipartite graph $\mathcal{G} = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{E})$ obtained by applying Proposition 1 and setting $w = q = k$. Thus, in the graph \mathcal{G} , $|\mathcal{V}_1| = |\mathcal{V}_2| = q^3 - q$. We use r to denote this number, that is $r = q^3 - q$. Recall that the degree of every vertex in \mathcal{G} is q .

Next, we modify this construction, to obtain a new graph $\mathcal{G}' = (\mathcal{V}'_1, \mathcal{V}'_2, \mathcal{E}')$ as follows. Let $\beta = \frac{1}{3} \frac{\epsilon}{4+\epsilon}$, and $b = \lceil (q^3 - q)^\beta \rceil$. To keep the presentation simple, from now on we assume that b divides q .

In the new graph \mathcal{G}' , the number of vertices on the right side remains the same as in \mathcal{G} , but we take b ‘‘copies’’ of the left hand side vertices. We distribute the edges adjacent to the left nodes of the graph between the ‘‘copies’’, thus the degree of every vertex in \mathcal{V}'_1 in \mathcal{G}' will be $k = \frac{q}{b}$, and the degree of every vertex in \mathcal{V}'_2 in \mathcal{G}' remains q . Note that $\beta < 1/3$ holds for any $\epsilon > 0$, thus $k > 1$ holds for large enough q .

More formally, let $|\mathcal{V}'_1| = b|\mathcal{V}_1| = br$, $\mathcal{V}'_2 = \mathcal{V}_2$, and $|\mathcal{E}'| = |\mathcal{E}|$. Thus each node in \mathcal{V}_1 corresponds to b nodes in \mathcal{V}'_1 and we think of these as copies of the original node in \mathcal{V}_1 . Let $\mathcal{V}'_1 = \{1, 2, \dots, br\}$. Then for $i' = 1, 2, \dots, br$, we have $i' = ur + i$ for $u = 0, 1, \dots, b-1$ and $i = 1, 2, \dots, r$. Thus, for a given node $i \in \mathcal{V}_1$, we have the copies $i, r+i, 2r+i, \dots, (b-1)r+i$ in \mathcal{V}'_1 .

We define the edges of \mathcal{G}' as follows. Suppose that the neighbors of node $i \in \mathcal{V}_1$ in \mathcal{G} are $\mathcal{N}(i) = \{j_1, j_2, \dots, j_q\}$. Then the neighbors of node $ur+i \in \mathcal{V}'_1$ in \mathcal{G}' are $\mathcal{N}'(ur+i) = \{j_{u\frac{q}{b}+1}, j_{u\frac{q}{b}+2}, \dots, j_{u\frac{q}{b}+\frac{q}{b}}\}$. That is, each copy of a given node i keeps a $\frac{1}{b}$ fraction of its original neighbors (and each vertex in \mathcal{V}'_1 has $k = q/b$ neighbors).

Next we prove that \mathcal{G}' has no 4-cycles or 6-cycles.

Assume that \mathcal{G}' has a 4-cycle, without loss of generality, suppose i'_1, i'_2, j_1, j_2 form a 4-cycle. Notice that if

$$i'_1 \equiv i'_2 \pmod{r} \tag{1}$$

then in graph \mathcal{G} , there are two edges between the nodes i_1 and j_1 , where $i_1 \equiv i'_1 \pmod{r}$, $1 \leq i_1 \leq r$. This is a contradiction, since \mathcal{G} does not have parallel edges. Otherwise, if (1) does not hold, then there exists a 4-cycle (formed by i_1, i_2, j_1, j_2) in the graph \mathcal{G} , where $i_1 \equiv i'_1 \pmod{r}$, $i_2 \equiv i'_2 \pmod{r}$, and $1 \leq i_1, i_2 \leq r$. This is a contradiction, since \mathcal{G} does not have 4-cycles.

Next, assume that \mathcal{G}' has a 6-cycle, without loss of generality, suppose $i'_1, i'_2, i'_3, j_1, j_2, j_3$ form a 6-cycle. Let us consider the following congruences:

$$\begin{aligned} i'_1 &\equiv i'_2 \pmod{r} \\ i'_2 &\equiv i'_3 \pmod{r} \\ i'_3 &\equiv i'_1 \pmod{r} \end{aligned} \tag{2}$$

If none of the congruences (2) holds, then there exists a 6-cycle in the graph \mathcal{G} . That 6-cycle is formed by $i_1, i_2, i_3, j_1, j_2, j_3$, where $i_t \equiv i'_t \pmod{r}$ and $1 \leq i_t \leq r$ for $t = 1, 2, 3$. This is a contradiction, since \mathcal{G} does not have 6-cycles. Otherwise, suppose that at least one of the congruences (2) holds. Without loss of generality, assume that $i'_1 \equiv i'_2 \pmod{r}$. Since $i'_1, i'_2, i'_3, j_1, j_2, j_3$ form a 6-cycle, one of j_1, j_2, j_3 , say j_1 is connected to both i'_1 and i'_2 . Then there are two edges between i_1 and j_1 in the graph \mathcal{G} , where $i_1 \equiv i'_1 \pmod{r}$ and $1 \leq i_1 \leq r$. This is a contradiction, since \mathcal{G} does not have parallel edges.

Thus, we have shown that \mathcal{G}' does not have 4-cycles or 6-cycles. Then, by Theorem 1 we obtain a primitive multiset batch code with number of servers $m = |\mathcal{V}'_1| + |\mathcal{V}'_2| = (b+1)r$, $k = q/b$, and rate $\frac{b}{b+1} = 1 - \frac{1}{b+1} > 1 - \frac{1}{b}$. Substituting $b = \lceil (q^3 - q)^\beta \rceil$ where $\beta = \frac{1}{3} \frac{\epsilon}{4+\epsilon}$, and expressing both m and the rate as a function of k we get

$$m = k^{3+\frac{12\beta}{1-3\beta}} + k^{3+\frac{9\beta}{1-3\beta}} = k^{3+\epsilon} + o(k^{3+\epsilon})$$

and rate at least

$$1 - \frac{1}{k^{\frac{3\beta}{1-3\beta}}} = 1 - \frac{1}{k^{\epsilon/4}}.$$

Applying the Gadget Lemma with $g = \lceil \frac{n}{k^{3+\epsilon}} \rceil$ we obtain an (n, N, k, m) multiset batch code with

$$N = \lceil \frac{n}{k^{3+\epsilon}} \rceil m$$

and rate at least $1 - \frac{1}{k^{\epsilon/4}}$. □

Theorem 4. *For any integer $c \geq 2$, any large enough prime power q , and large enough integer n , there is an explicit (n, N, k, m) multiset batch code with rate $1 - \frac{1}{c}$, $m \leq c(c-1)^3 k^3 = O(k^3)$ and*

$$N \leq \left\lceil \frac{n}{(c-1)(q^3 - q)} \right\rceil m \leq \left\lceil \frac{1.1n}{(c-1)^4 k^3} \right\rceil m = O\left(\frac{n}{k^3} m\right)$$

where $k = \lfloor \frac{q}{b} \rfloor$ and $b = c - 1$.

Proof. Substitute $k = \lfloor \frac{q}{b} \rfloor$ and $b = c - 1$ in the above construction. □

4.2 Constructions based on the Lazebnik, Ustimenko, Woldar graphs [20]

Here, we present the construction of bipartite graphs with girth at least 8 from [20]. These graphs allow for a left degree k which is polynomial in the number of left nodes n while achieving the rate $1 - o_k(1)$. First, we briefly describe the construction of these bipartite graphs and then comment on the parameters of the multiset batch codes obtained based on these graphs.

Lazebnik et al. construction : Given an odd prime power q , let $t \in (0, 2]$ and $s \in (0, 1]$ be such that q^t and q^s are integers. The bipartite graph $\mathcal{B}_{s,t}(q) = \mathcal{G} = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{E})$ with q^{3+t} left nodes and q^{3+s} right nodes is constructed as follows. Let $\mathcal{T} \subseteq \mathbb{F}_{q^2}$ and $\mathcal{S} \subseteq \mathbb{F}_q$ with $|\mathcal{T}| = q^t$ and $|\mathcal{S}| = q^s$. Associate q^{3+t} left nodes with 3-dimensional vectors such that

$$\mathcal{V}_1 = \{\mathbf{l} = (l_1, l_2, l_3) : l_1 \in \mathcal{T} \subseteq \mathbb{F}_{q^2}, l_2 \in \mathbb{F}_{q^2}, l_3 \in \mathbb{F}_q\}.$$

The q^{3+s} right nodes of the bipartite graph \mathcal{G} are denoted by 3-dimensional vectors such that

$$\mathcal{V}_2 = \{\mathbf{v} = (v_1, v_2, v_3) : v_1 \in \mathcal{S} \subseteq \mathbb{F}_q, v_2 \in \mathbb{F}_{q^2}, v_3 \in \mathbb{F}_q\}.$$

In the graph \mathcal{G} , there exists an edge between the left node $\mathbf{l} = (l_1, l_2, l_3)$ and the right node $\mathbf{v} = (v_1, v_2, v_3)$, i.e., $(\mathbf{l}, \mathbf{v}) \in \mathcal{E}$, iff we the following two conditions are satisfied

$$l_2 - v_2 = l_1 v_1, \tag{3a}$$

$$l_3 - v_3 = f(l_1)v_2 + l_1 f(v_2). \tag{3b}$$

Here, $f : x \mapsto f(x)$ represents the involutive automorphism of \mathbb{F}_{q^2} with fixed field \mathbb{F}_q . Lazebnik *et al.* establish that the graph $\mathcal{B}_{s,t}(q)$ has girth at least 8 [20, Theorem 3]. Here, we paraphrase the result to make it consistent with the rest of the paper.

Proposition 2. [20, Theorem 3] *Let q be an odd prime power. Then for any $t \in (0, 2]$ and $s \in (0, 1]$ such that both q^t and q^s are integers, the bipartite graph $\mathcal{B}_{s,t}(q)$ is biregular with $q^{3+s} + q^{3+t}$ total nodes (q^{3+t} left nodes, and q^{3+s} right nodes), q^{3+s+t} edges and girth at least 8. The left degree and the right degree of $\mathcal{B}_{s,t}(q)$ are q^s and q^t , respectively.*

Remark 4. For the case when $\mathcal{T} = \mathbb{F}_{q^2}$ and $q \geq 3$, Lazebnik *et al.* show that the graph $\mathcal{B}_s(q) = \mathcal{B}_{s,2}(q)$ has girth exactly equal to 8.

We now present the main result of this subsection, which states the parameters of the multiset batch codes obtained from the bipartite graphs in [20].

Theorem 5. *For any $0 < \epsilon \leq 1$, one can find an infinite sequence of odd prime powers k such that for large enough integer n , there is an explicit (n, N, k, m) multiset batch code with rate at least $1 - \frac{1}{k^\epsilon}$, $m = k^{4+\tilde{\epsilon}} + o(k^{4+\tilde{\epsilon}})$, and*

$$N = \left\lceil \frac{n}{k^{4+\tilde{\epsilon}}} \right\rceil m.$$

Here, $\tilde{\epsilon}$ is a rational number which can be taken arbitrarily close to ϵ .

Proof. From Proposition 2, we know that for any $t \in (1, 2]$ such that q^t is an integer, the bipartite graph $\mathcal{G} = \mathcal{B}_{s=1,t}(q)$ obtained by Lazebnik *et al.* construction has q^{3+t} left nodes, $q^{3+s} = q^4$ right nodes, and girth at least 8. Also, note that each left node in $\mathcal{B}_{s=1,t}(q)$ has degree $q^s = q$. Therefore, it follows from Theorem 1 that the bipartite graph $\mathcal{B}_{s=1,t}(q)$ gives an $(n_1 = q^{3+t}, N_1 = q^{3+t} + q^4, k = q, m = N_1)$ primitive batch code \mathcal{C} . The rate of \mathcal{C} is $\frac{q^{3+t}}{q^{3+t} + q^4} \geq 1 - \frac{q^4}{q^{3+t}} = 1 - \frac{1}{q^{t-1}}$.

Now, for an $\epsilon > 0$, we pick a rational number $\tilde{\epsilon}$ such that $\epsilon \leq \tilde{\epsilon}$ for $\tilde{\epsilon}$ arbitrarily close to ϵ . By setting $t = 1 + \tilde{\epsilon}$ and choosing $k = q$ to be an odd prime power so that $q^{1+\tilde{\epsilon}}$ is an integer⁷, we have that \mathcal{C} is an $(n_1 = k^{4+\tilde{\epsilon}}, N_1 = k^{4+\tilde{\epsilon}} + k^4, k, m = N_1)$ primitive batch code with rate at least $1 - \frac{1}{k^\epsilon} \geq 1 - \frac{1}{k^\epsilon}$. We now apply the gadget lemma (see Section 2) with $g = \left\lceil \frac{n}{n_1} \right\rceil = \left\lceil \frac{n}{k^{4+\tilde{\epsilon}}} \right\rceil$ to obtain an (n, N, k, m) batch code with rate at least $1 - \frac{1}{k^\epsilon}$, $m = k^{4+\tilde{\epsilon}} + o(k^{4+\tilde{\epsilon}})$, and

$$N = \left\lceil \frac{n}{k^{4+\tilde{\epsilon}}} \right\rceil m.$$

□

4.3 Constructions based on the de Caen, Székely graphs

In [7], de Caen and Székely used known constructions of generalized quadrangles to construct bipartite graphs with maximum number of edges and no 4- and 6-cycles. Generalized quadrangles are combinatorial structures defined as follows. Let \mathcal{P} be a set of *points*, and let \mathcal{L} be a collection of subsets of \mathcal{P} called *lines*. $(\mathcal{P}, \mathcal{L})$ is called a *generalized quadrangle* of order (s, t) ($s, t \geq 1$) if the following holds:

1. Each point is incident with $t + 1$ lines, and two distinct points are incident with at most one line.
2. Each line is incident with $s + 1$ points, and two distinct lines are incident with at most one point.

⁷Note that for any rational number $\tilde{\epsilon}$ there is an infinite sequence of such odd prime powers.

3. If $x \in V$ is a point and $L \in \mathcal{L}$ is a line not incident with x , then there is a unique pair (y, M) , where M is a line and y is a point incident to M , such that x is incident to M and y is incident to L .

The conditions above imply that the number of points $|\mathcal{P}| = (s+1)(st+1)$ and the number of lines $|\mathcal{L}| = (t+1)(st+1)$. Given any generalized quadrangle, we can represent it by a bipartite graph $\mathcal{G} = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{E})$ in the straightforward way, by letting $\mathcal{V}_1 = \mathcal{L}$, $\mathcal{V}_2 = \mathcal{P}$ and $(L, x) \in \mathcal{E}$ if and only if x is incident to L . Note that the number of edges of this graph is $|\mathcal{L}|(s+1) = |\mathcal{P}|(t+1) = (s+1)(t+1)(st+1)$.

The first two requirements in the definition of generalized quadrangles immediately imply that the bipartite graph representing a generalized quadrangle cannot contain any 4-cycles, and the last condition immediately implies that the graph cannot contain any 6-cycles. De Caen and Székely observed that the bipartite graphs representing generalized quadrangles in fact contain the largest possible number of edges among graphs on the same number of vertices without 4- and 6-cycles.

Generalized quadrangles have been extensively studied and several constructions are known for settings where s and t are close to each other, e.g when $s = t$ or $s = q-1$ and $t = q+1$. However, to obtain multiset batch codes with rate close to 1, we need generalized quadrangles where s and t have different orders of magnitude. There are two families of constructions of generalized quadrangles of this type known: for any prime power q , there are generalized quadrangles of order (q, q^2) and of order (q^2, q^3) . For more details on generalized quadrangles see [27, 40].

Generalized quadrangles of order (q, q^2) yield multiset batch codes with $m = (1 + o_k(1))k^5$ and rate $1 - 1/(k-1)$. The performance of these codes is similar to the codes based on the Lazebnik et al. construction when taking $\epsilon = 1$ in that construction.

Generalized quadrangles of order (q^2, q^3) yield multiset batch codes with $m = (1 + o_k(1))k^4$ and rate $1 - 1/\sqrt{k}$.

Theorem 6. *Let q be any prime power, and let $k = q^2$. Then for any large enough integer n there is an explicit (n, N, k, m) multiset batch code with rate at least $1 - 1/\sqrt{k}$, $m = (1 + o_k(1))k^4$, and $N = \lceil \frac{n}{k^4} \rceil m$.*

Proof. The known constructions of generalized quadrangles of order (q^2, q^3) immediately give bipartite graphs with $n_1 = |\mathcal{V}_1| = (q^3+1)(q^5+1)$, $|\mathcal{V}_2| = (q^2+1)(q^5+1)$, and left degree $k = q^2+1$. We also know that these bipartite graphs have no 4- and 6-cycles. Using Theorem 1, this gives (n_1, N_1, k, m) primitive multiset batch codes with $N_1 = m = |\mathcal{V}_1| + |\mathcal{V}_2| = q^8 + o_k(q^8) = (1 + o_k(1))k^4$. The rate of the resulting primitive multiset batch code is $|\mathcal{V}_1|/(|\mathcal{V}_1| + |\mathcal{V}_2|) \geq 1 - 1/q = 1 - 1/\sqrt{k}$. Applying the gadget lemma as in the previous subsection gives the statement of the theorem. \square

4.4 Constructions based on zig-zag codes of [39]

In this subsection, we describe a construction for primitive multiset batch codes that allow for $m = \Theta(k^{k^{1+\epsilon}+1+\epsilon})$ with any $0 < \epsilon$ while achieving rate at least $1 - (1/k^\epsilon)$. In particular, we present a construction of biregular bipartite graphs which are free of 4- and 6-cycles. For a given prime k and an integer r , the graph has rk^r left nodes and k^{r+1} right nodes. The left degree and right degree of these graphs are k and r , respectively. The construction presented here is based on the work of Tamo *et al.* on zig-zag codes [39]. This construction has been previously employed in [31], where it gives codes with disjoint repair groups, *i.e.*, the bipartite graph is free of 4-cycles. Here, we go one step further and show that the bipartite graphs obtained from zig-zag construction are free of 6-cycles as well. It then follows from Theorem 1 that these bipartite graphs give multiset batch codes.

Zig-zag construction : Let k be a prime number⁸ and r be an integer. Each of the rk^r left nodes of the bipartite graph are represented by an $(r + 1)$ -dimensional vectors such that

$$\mathcal{V}_1 = \{\mathbf{l} = (l_0, l_1, l_2, \dots, l_r) : l_0 \in [r] \text{ and } l_i \in [0 : (k - 1)] \text{ for } i \in [r]\}.$$

Similarly, k^{r+1} right nodes of the bipartite graphs are denoted by an $(r + 1)$ -dimensional vectors such that

$$\mathcal{V}_2 = \{\mathbf{v} = (v_0, v_1, v_2, \dots, v_r) : v_i \in [0 : (k - 1)] \text{ for } i \in [0 : r]\}.$$

Given a left node $\mathbf{l} = (l_0, l_1, l_2, \dots, l_r)$ and a right node $\mathbf{v} = (v_0, v_1, v_2, \dots, v_r)$, there is an edge between \mathbf{l} and \mathbf{v} , *i.e.*, $(\mathbf{l}, \mathbf{v}) \in \mathcal{E}$ iff

$$(l_1, l_2, \dots, l_r) + v_0 \mathbf{e}_{l_0} = (v_1, v_2, \dots, v_r) \pmod{k}. \quad (4)$$

Here, $\mathbf{e}_{l_0} \in \mathbb{Z}^r$ denotes the standard unit vector which has 1 at the l_0 -th coordinate and zeros at other $r - 1$ coordinates. From (4), a given left node \mathbf{l} , has k neighbors for k values of v_0 . Similarly, a right node \mathbf{v} has r left neighbors as for each $l_0 \in [r]$ there is a unique vector $(l_1, l_2, \dots, l_r) \in [0 : (k - 1)]^r$ such that (4) holds.

In [31, Claim 1], it is shown that the bipartite graph obtained using zig-zag construction has no 4-cycle⁹. Here, we state this result without its proof.

Lemma 4. [31] *The biregular bipartite graph obtained by zig-zag construction has no 4-cycle.*

Next, we establish that the bipartite graphs generated by zig-zag construction are free of 6-cycles as well.

Lemma 5. *Let \mathcal{G} be a biregular bipartite graph obtained from zig-zag construction. Then, the girth of \mathcal{G} is at least 8.*

Proof. We have from Lemma 4 that \mathcal{G} is free of 4-cycles. Here, we show that it is not possible to have a 6-cycle in \mathcal{G} which establishes the Lemma 5. In order to show contradiction, we assume that there is a 6-cycle $\{\mathbf{l}^1, \mathbf{v}^1, \mathbf{l}^2, \mathbf{v}^2, \mathbf{l}^3, \mathbf{v}^3\}$ in the graph \mathcal{G} . Here, $\{\mathbf{l}^1, \mathbf{l}^2, \mathbf{l}^3\}$ and $\{\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3\}$ represent three distinct left and right nodes, respectively. From (4) we have the following relations.

$$(l_1^1, l_2^1, \dots, l_r^1) + v_0^1 \mathbf{e}_{l_0^1} = (v_1^1, v_2^1, \dots, v_r^1) \pmod{k}, \quad (5a)$$

$$(l_1^2, l_2^2, \dots, l_r^2) + v_0^1 \mathbf{e}_{l_0^2} = (v_1^1, v_2^1, \dots, v_r^1) \pmod{k}. \quad (5b)$$

$$(l_1^2, l_2^2, \dots, l_r^2) + v_0^2 \mathbf{e}_{l_0^2} = (v_1^2, v_2^2, \dots, v_r^2) \pmod{k}, \quad (6a)$$

$$(l_1^3, l_2^3, \dots, l_r^3) + v_0^2 \mathbf{e}_{l_0^3} = (v_1^2, v_2^2, \dots, v_r^2) \pmod{k}. \quad (6b)$$

and

$$(l_1^3, l_2^3, \dots, l_r^3) + v_0^3 \mathbf{e}_{l_0^3} = (v_1^3, v_2^3, \dots, v_r^3) \pmod{k}, \quad (7a)$$

$$(l_1^1, l_2^1, \dots, l_r^1) + v_0^3 \mathbf{e}_{l_0^1} = (v_1^3, v_2^3, \dots, v_r^3) \pmod{k}. \quad (7b)$$

⁸The construction also works when k is a prime power. However, for the ease of exposition, we only present the construction when k is a prime number.

⁹The Claim 1 in [31] states this result in an alternative manner.

By subtracting (7b) from (5a) we obtain

$$(v_0^1 - v_0^3)\mathbf{e}_{l_0^1} = (v_1^1 - v_1^3, v_2^1 - v_2^3, \dots, v_r^1 - v_r^3) \pmod{k}. \quad (8)$$

We need to have $v_0^1 \neq v_0^3$, otherwise from (8) we have $\mathbf{v}^1 = \mathbf{v}^2$, which is not true as \mathbf{v}^1 and \mathbf{v}^2 are distinct nodes. By subtracting (5b) from (5a), we obtain

$$(l_1^1 - l_1^2, l_2^1 - l_2^2, \dots, l_r^1 - l_r^2) + v_0^1(\mathbf{e}_{l_0^1} - \mathbf{e}_{l_0^2}) = 0. \quad (9)$$

Again, we need to have $l_0^1 \neq l_0^2$. Otherwise, (9) implies that $\mathbf{l}^1 = \mathbf{l}^2$, which is not true. Similarly, we can obtain the following sets of equations

$$(v_0^1 - v_0^2)\mathbf{e}_{l_0^2} = (v_1^1 - v_1^2, v_2^1 - v_2^2, \dots, v_r^1 - v_r^2) \pmod{k}, \quad (10)$$

$$(v_0^2 - v_0^3)\mathbf{e}_{l_0^3} = (v_1^2 - v_1^3, v_2^2 - v_2^3, \dots, v_r^2 - v_r^3) \pmod{k}. \quad (11)$$

and

$$(l_1^2 - l_1^3, l_2^2 - l_2^3, \dots, l_r^2 - l_r^3) + v_0^2(\mathbf{e}_{l_0^2} - \mathbf{e}_{l_0^3}) = 0, \quad (12)$$

$$(l_1^3 - l_1^1, l_2^3 - l_2^1, \dots, l_r^3 - l_r^1) + v_0^3(\mathbf{e}_{l_0^3} - \mathbf{e}_{l_0^1}) = 0. \quad (13)$$

Combining (9), (12) and (13), we obtain that l_0^1, l_0^2 and l_0^3 are three distinct integers in $[r]$. Similarly, it follows from (8), (10), and (11) that v_0^1, v_0^2 and v_0^3 are distinct integers. Moreover, we have from (8), (10), and (11) that

$$v_{l_0^1}^1 \neq v_{l_0^1}^3 \quad (14a)$$

$$v_z^1 = v_z^3, \text{ for } z \in [r] \setminus \{l_0^1\}. \quad (14b)$$

$$v_{l_0^2}^1 \neq v_{l_0^2}^2 \quad (15a)$$

$$v_z^1 = v_z^2, \text{ for } z \in [r] \setminus \{l_0^2\}. \quad (15b)$$

and

$$v_{l_0^3}^2 \neq v_{l_0^3}^3 \quad (16a)$$

$$v_z^2 = v_z^3, \text{ for } z \in [r] \setminus \{l_0^3\}. \quad (16b)$$

Note that l_0^1, l_0^2 and l_0^3 are distinct integers. Therefore, the equation (14a) to (16b) lead to contradiction as it is not possible to simultaneously satisfy all of these equations. Thus, there is no 6-cycle in the graph \mathcal{G} . \square

The following result presents the parameters of the multiset batch codes obtained from zig-zag construction .

Theorem 7. *For any fixed integer k , a large enough integer n , and $\epsilon > 0$, there is an explicit (n, N, k, m) multiset batch code with rate at least $1 - \frac{1}{k^\epsilon}$, $m = \Theta(k^{k^{1+\epsilon}+1+\epsilon})$, and*

$$N = \left\lceil \frac{n}{k^{k^{1+\epsilon}+1+\epsilon}} \right\rceil m.$$

Proof. From Lemma 5, we know that a bipartite graph \mathcal{G} obtained by zig-zag construction has rk^r left nodes and k^{r+1} right nodes, and girth at least 8. Moreover, the degree of each left node in \mathcal{G} is k . Therefore, it follows from Theorem 1 that \mathcal{G} provides an $(n_1 = rk^r, N_1 = (r+k)k^r, k, m = N_1)$ primitive batch code \mathcal{C} with rate $\frac{rk^r}{rk^r+k^{r+1}} = \frac{r}{r+k} \geq 1 - \frac{k}{r}$.

We now select $r = k^{1+\epsilon}$, which gives us the rate of the code at least $1 - \frac{1}{k^\epsilon}$ ¹⁰. This particular choice for the parameter r implies that \mathcal{C} is an $(n_1 = k^{k^{1+\epsilon}+1+\epsilon}, N_1 = (k^\epsilon + 1)k^{k^{1+\epsilon}+1}, k, m = N_1)$ primitive batch code. Now using gadget lemma (see Section 2) with $g = \left\lceil \frac{n}{n_1} \right\rceil = \left\lceil \frac{n}{k^{k^{1+\epsilon}+1+\epsilon}} \right\rceil$, we obtain an (n, N, k, m) batch code $g \cdot \mathcal{C}$ with rate at least $1 - \frac{1}{k^\epsilon}$, $m = \Theta(k^{k^{1+\epsilon}+1+\epsilon})$, and

$$N = \left\lceil \frac{n}{k^{k^{1+\epsilon}+1+\epsilon}} \right\rceil m.$$

□

4.5 Limitations of the bipartite graph based constructions

We note that the bipartite graph based constructions described above exhibit a tradeoff between the parameter m as a function of k , and the rate of the codes obtained as a function of k . Ideally, we would like to minimize m (the number of servers) as a function of k (the number of read requests that can be supported) and maximize the rate as a function of k .

Our constructions achieve close to optimal tradeoffs by the following estimates of de Caen and Székely [7].

Theorem 8. [7, 8] *Let $G = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{E})$ be a bipartite graph without 4- and 6-cycles, where $|\mathcal{V}_1| = n_1$, $|\mathcal{V}_2| = n_2$ and $\sqrt{n_1} \leq n_2 \leq n_1$. Then $|\mathcal{E}| = O(n_1^{2/3} n_2^{2/3})$.*

Note that we can assume $n_2 \geq \sqrt{n_1}$, since otherwise $|\mathcal{E}| = O(n_1)$ [8] which would only allow a constant number of read requests k .

These estimates imply that assuming rate at least $1/2$ the number of servers m has to be at least $\Omega(k^3)$ in multiset batch codes based on bipartite graphs without 4- and 6-cycles. More precisely, assuming rate at least $1 - \frac{1}{k^\alpha}$ implies $m = \Omega(k^{3+2\alpha})$. Thus, our results give close to optimal tradeoffs for multiset batch codes based on bipartite graphs.

We also note that any systematic primitive multiset batch code where decoding uses at most one non-systematic code symbol for every request (as in our code constructions) must have an associated bipartite graph without 4- and 6-cycles. Thus, our results in Sections 4.1 and 4.3 give close to optimal tradeoffs between the number of servers and the rate for multiset batch codes of this type.

5 On Fault Tolerance of Batch Codes

In [17], Ishai *et al.* study batch codes with *load balancing* as the primary focus of their work. However, in practice the codes for distributed storage systems need to allow for mechanisms to protect data against catastrophic events. These catastrophic events may correspond to the permanent or persistent loss of a large number of servers due to multiple reasons, including hardware failures, power outages, and loss of rack switches in data centers [9, 12]. Therefore, it is desirable to work with codes that have large minimum

¹⁰If $k^{1+\epsilon}$ is not an integer, r can be made integer by taking the floor or the ceiling of $k^{1+\epsilon}$.

distance. Here, we note that Lipmaa and Skachek have recently shown that a binary linear batch code which supports any k -request patterns has minimum distance at least k [22].

Here, we describe a straightforward yet powerful way to combine the notion of load balancing in batch codes with the issue of fault tolerance. We show that one can modify a given batch code to get a code with large minimum distance at the cost of rate loss. In particular, given a field of large enough size and a batch code with rate $1 - o_k(1)$, one can get a code which is also a batch code and has minimum distance arbitrarily close to the Singleton bound.

Lemma 6. *Let \mathcal{C}^G be an $[n, N^G, d^G]$ ¹¹ systematic linear code with minimum distance d^G , and \mathcal{C}^B be an $(n, N^B, k, m^B = N^B)$ systematic primitive batch code. Then, it is possible to obtain an $(n, N = N^B + N^G - n, k, m = N)$ systematic primitive batch code $\mathcal{C} = \mathcal{C}^G \circ \mathcal{C}^B$ with distance at least d^G .*

Proof. Let \mathcal{C}^G be an $[n, N^G = n + p, d^G]$ systematic linear code with minimum distance d^G which encodes a length- n message vector $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \Sigma^n$ to a codeword

$$\mathbf{c}^G = (c_1^G = x_1, c_2^G = x_2, \dots, c_n^G = x_n, c_{n+1}^G, \dots, c_{N^G}^G) \in \Sigma^{N^G}.$$

Let \mathcal{C}^B be an $(n, N^B, k, m = N^B)$ systematic primitive batch code which encodes the length- n message vector $(x_1, x_2, \dots, x_n) \in \Sigma^n$ to a codeword

$$\mathbf{c}^B = (c_1^B = x_1, c_2^B = x_2, \dots, c_n^B = x_n, c_{n+1}^B, \dots, c_{N^B}^B).$$

Given \mathcal{C}^B and \mathcal{C}^G , we obtain a systematic code $\mathcal{C} = \mathcal{C}^G \circ \mathcal{C}^B$ which encodes the length- n message vector $(x_1, x_2, \dots, x_n) \in \Sigma^n$ to a codeword

$$\mathbf{c} = (c_1, c_2, \dots, c_{N=N^B+p}) = (x_1, x_2, \dots, x_n, c_{n+1}^B, \dots, c_{N^B}^B, c_{n+1}^G, \dots, c_{N^G}^G) \in \Sigma^{N^B+N^G-n}.$$

It is clear from the construction that the minimum distance of \mathcal{C} is at least as large as the minimum distance of \mathcal{C}^G as the codewords of \mathcal{C}^G are sub-vectors of the codewords of \mathcal{C} . Similarly, one can obtain the codeword corresponding to \mathbf{x} in \mathcal{C}^B by puncturing the codeword associated with \mathbf{x} in \mathcal{C} . Therefore, \mathcal{C} is a batch code that can support sequences of k read requests. □

Remark 5. Note that Lemma 6 holds for multiset batch codes as well. Also, we state Lemma 6 only for primitive batch codes. One can employ the gadget lemma (see Section 2) with primitive multiset batch code $\mathcal{C} = \mathcal{C}^G \circ \mathcal{C}^B$ to obtain general multiset batch codes which can tolerate the failure of any $d^G - 1$ servers (buckets).

Taking Σ to be a field with large enough size and \mathcal{C}^G to a systematic MDS code, we obtain the following result. Note that if Σ is not sufficiently large, we can create parities over the extension field Σ^B .

Lemma 7. *For large enough integers k and n , let \mathcal{C}^B be an $(n, N^B, k, m^B = N^B)$ systematic primitive batch code with rate at least $1 - o_k(1)$. Further assume that Σ is a field with large enough size. Then for any rate $0 < \rho < 1$, it is possible to obtain an $(n, N, k, m = N)$ primitive batch code \mathcal{C} with minimum distance $d \geq (N - n + 1) - o_k(n)$.*

Proof. From the assumption on the rate of \mathcal{C}^B , we know that

$$n \geq (1 - o_k(1))N^B. \tag{17}$$

¹¹An $[n, N, d]$ code is an $[n, N]$ code with minimum distance d .

Now, we take \mathcal{C}^G in Lemma 6 to be an $[n, N^G, N^G - n + 1]$ systematic MDS code, which gives us an $(n, N = N^B + N^G - n, k, m = N)$ systematic primitive batch code with minimum distance

$$d \geq N^G - n + 1.$$

The above expression can be rewritten as

$$\begin{aligned} d &\geq N + n - N^B - n + 1 \\ &= N - n + 1 - (N^B - n) \\ &\stackrel{(i)}{\geq} N - n + 1 - o_k(1)N^B \\ &\stackrel{(ii)}{=} N - n + 1 - o_k(n). \end{aligned} \tag{18}$$

Here, (i) and (ii) follow from (17). The rate ρ of \mathcal{C} is equal to $\frac{\rho^G}{1+o_k(1)}$, where $\rho^G = \frac{n}{N^G}$ denotes the rate of \mathcal{C}^G . Note that we can select ρ^G to be any number in the interval $(0, 1)$. \square

References

- [1] N. Balachandran and S. Bhattacharya. On an extremal hypergraph problem related to combinatorial batch codes. *Discrete Applied Mathematics*, 162:373–380, 2014.
- [2] C. Balbuena. Incidence matrices of projective planes and other bipartite graphs of few vertices. *SIAM Journal of Discrete Maths*, 22(4):1351–1363, 2008.
- [3] C. Balbuena. A construction of small regular bipartite graphs of girth 8. *Discrete Mathematics and Theoretical Computer Science*, 11(2):33–46, 2009.
- [4] S. Bhattacharya, S. Ruj, and B. Roy. Combinatorial batch codes: A lower bound and optimal constructions. *Advances in Mathematics of Communications*, 6(2), 2012.
- [5] R. A. Brualdi, K. P. Kiernan, S. A. Meyer, and M. W. Schroeder. Combinatorial batch codes and transversal matroids. *Adv. Math. Commun.*, 4(3):419–431, 2010.
- [6] C. Bujtás and Z. Tuza. Optimal combinatorial batch codes derived from dual systems. *Miskolc Mathematical Notes*, 12(1), 2011.
- [7] D. de Caen and L. A. Székely. The maximum size of 4- and 6-cycle free bipartite graphs on m, n vertices. In *Graphs, Sets and Numbers (Proc. of the Conference Dedicated to the 60th Birthdays of A. Hajnal and Vera T. Sós)*, Coll. Math. Soc. J. Bolyai, pages 135–142, 1991.
- [8] D. de Caen and L. A. Székely. On dense bipartite graphs of girth eight and upper bounds for certain configurations in planar point-line systems. *Journal of Combinatorial Theory*, 77:268–278, 1997.
- [9] J. Dean. Evolution and future directions of large-scale storage and computation systems at google. <http://research.google.com/people/jeff/SOCC2010-keynote.html>, 2010.
- [10] A. G. Dimakis, P. Godfrey, Y. Wu, M. Wainwright, and K. Ramchandran. Network coding for distributed storage systems. *IEEE Trans. Inf. Theory*, 56(9):4539–4551, 2010.

- [11] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh. A survey on network codes for distributed storage. *Proc. of the IEEE*, 99(3):476–489, 2011.
- [12] D. Ford, F. Labelle, F. I. Popovici, M. Stokely, V. Truong, L. Barroso, C. Grimes, and S. Quinlan. Availability in globally distributed storage systems. In *Presented as part of the 9th USENIX Symposium on Operating Systems Design and Implementation*, Berkeley, CA, 2010. USENIX.
- [13] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin. On the locality of codeword symbols. *IEEE Trans. Inf. Theory*, 58(11):6925–6934, 2012.
- [14] A. Guo, S. Kopparty, and M. Sudan. New affine-invariant codes from lifting. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 529–540. ACM, 2013.
- [15] B. Hemenway, R. Ostrovsky, and M. Wootters. Local correctability of expander codes. In *Automata, Languages, and Programming*, pages 540–551. Springer, 2013.
- [16] C. Huang, H. Simitci, Y. Xu, A. Ogus, B. Calder, P. Gopalan, J. Li, S. Yekhanin, et al. Erasure coding in windows azure storage. In *USENIX Annual Technical Conference (ATC)*, pages 15–26, 2012.
- [17] Y. Ishai, E. Kushilevitz, R. Ostrovsky, and A. Sahai. Batch codes and their applications. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, STOC '04, pages 262–271, New York, NY, USA, 2004. ACM.
- [18] G. Kamath, N. Prakash, V. Lalitha, and P. Kumar. Codes with local regeneration. *CoRR*, abs/1211.1932, 2012.
- [19] S. Kopparty, S. Saraf, and S. Yekhanin. High-rate codes with sublinear-time decoding. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 167–176. ACM, 2011.
- [20] F. Lazebnik, V. A. Ustimenko, and A. J. Woldar. New constructions of bipartite graphs on m, n vertices with many edges and without small cycles. *Journal of Combinatorial Theory*, 58:111–117, 1994.
- [21] F. Lazebnik, V. A. Ustimenko, and A. J. Woldar. A new series of dense graphs of high girth. *Bulletin of the American Mathematical Society*, 32(1):73–79, 1995.
- [22] H. Lipmaa and V. Skachek. Linear batch codes. *CoRR*, abs/1404.2796, 2014.
- [23] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland, 1983.
- [24] F. Oggier and A. Datta. Self-repairing codes for distributed storage - a projective geometric construction. In *Proc. of ITW*, 2011.
- [25] D. Papailiopoulos and A. G. Dimakis. Locally repairable codes. In *Proc. of IEEE ISIT*, 2012.
- [26] M. B. Paterson, D. R. Stinson, and R. Wei. Combinatorial batch codes. *Advances in Mathematics of Communications*, 3(1):13–27, 2009.
- [27] S. E. Payne and J. A. Thas. Finite generalized quadrangles. *Research Notes in Mathematics, 110*. Pitman (Advanced Publishing Program), Boston, MA, 1984.

- [28] J. S. Plank et al. A tutorial on reed-solomon coding for fault-tolerance in raid-like systems. *Softw., Pract. Exper.*, 27(9):995–1012, 1997.
- [29] M. O. Rabin. Efficient dispersal of information for security, load balancing, and fault tolerance. *Journal of the ACM (JACM)*, 36(2):335–348, 1989.
- [30] A. S. Rawat, O. O. Koyluoglu, N. Silberstein, and S. Vishwanath. Optimal locally repairable and secure codes for distributed storage systems. *IEEE Trans. Inf. Theory*, 60(1):212–236, 2014.
- [31] A. S. Rawat, D. S. Papailiopoulos, A. G. Dimakis, and S. Vishwanath. Locality and availability in distributed storage. In *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, pages 681–685, 2014.
- [32] I. S. Reed and G. Solomon. Polynomial codes over certain finite fields. *Journal of the Society for Industrial & Applied Mathematics*, 8(2):300–304, 1960.
- [33] T. Richardson and R. Urbanke. *Modern coding theory*. Cambridge University Press, New York, NY, USA, 2008.
- [34] S. Ruj and B. K. Roy. More on combinatorial batch codes. *CoRR*, abs/0809.3357, 2008.
- [35] M. Sathiamoorthy, M. Asteris, D. Papailiopoulos, A. G. Dimakis, R. Vadali, S. Chen, and D. Borthakur. Xoring elephants: Novel erasure codes for big data. In *Proceedings of the 39th international conference on Very Large Data Bases*, pages 325–336. VLDB Endowment, 2013.
- [36] N. Silberstein and A. Gál. Optimal combinatorial batch codes based on block designs. *CoRR*, abs/1312.5505, 2013.
- [37] I. Tamo and A. Barg. A family of optimal locally recoverable codes. *CoRR*, abs/1311.3284, 2013.
- [38] I. Tamo, D. Papailiopoulos, and A. G. Dimakis. Optimal locally repairable codes and connections to matroid theory. In *Proc. of IEEE ISIT*, 2013.
- [39] I. Tamo, Z. Wang, and J. Bruck. Zigzag codes: Mds array codes with optimal rebuilding. *IEEE Trans. Inf. Theory*, 59(3):1597–1616, 2013.
- [40] K. Thas. *Symmetry in Finite Generalized Quadrangles*. Frontiers in Mathematics. Springer Verlag NY, 2004.
- [41] V. Ustimenko. A linear interpretation of the flag geometries of chevalley groups. *Ukrainian Mathematical Journal*, 42(3):341–344, 1990.
- [42] S. Yekhanin. *Locally decodable codes and private information retrieval schemes*. Springer, 2010.