High-rate storage codes on triangle-free graphs

Alexander Barg* Gilles Zémor†

January 30, 2022

Abstract

Consider an assignment of bits to the vertices of a connected graph $G(V, E)$ with the property that the value of each vertex is a function of the values of its neighbors. A collection of such assignments is called a *storage code* of length $|V|$ on G. The storage code problem can be equivalently formulated as maximizing the probability of success in a *guessing game* on graphs, or constructing *index codes* of small rate.

If G contains many cliques, it is easy to construct codes of rate close to 1, so a natural problem is to construct high-rate codes on triangle-free graphs, where constructing codes of rate $> 1/2$ is a nontrivial task, with few known results. In this work we construct infinite families of linear storage codes with high rate relying on coset graphs of binary linear codes. We also derive necessary conditions for such codes to have high rate, and even rate potentially close to 1.

Extended abstract[1]

1 Context

In this paper we consider a class of codes on graphs known as storage codes. Given an undirected graph $G(V, E)$ with N vertices, denote by $\mathcal{N}(v) = \{u : (v, u) \in E\}$ the set of neighbors of the vertex v in G. Consider a set of vectors $\mathcal{C} = \{x : x \in Q^N\}$ over a finite alphabet Q, where the coordinates are indexed by the vertices in V. The set $\mathcal{C}$ is said to form a *storage code* if for every $v \in V$ and $x, y \in \mathcal{C}$, if $x_u = y_u$ for all $u \in \mathcal{N}(v)$ then also $x_v = y_v$. In other words, the codewords are written on the vertices of G, and for any codeword the value of the vertex v can be uniquely determined by the values of its neighbors. Thinking of storing the coordinates of the codeword at different nodes of a distributed storage system, this definition implies that an erased coordinate (vertex) can be recovered in a local way from its immediate neighborhood.

The concept of storage codes on graphs was introduced around 2014 by Mazumdar [14, 15] and Shanmugam and Dimakis [17], motivated by earlier works on index coding by Alon et al. [1] and codes with locality for distributed storage introduced by Gopalan et al. [12]. The related

*Department of ECE and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA and Inst. for Probl. Inform. Trans., Moscow, Russia abarg@umd.edu

†Mathematics Institute, UMR 5251, University of Bordeaux, France. zemor@math.u-bordeaux.fr

[1]Full paper available at https://arxiv.org/abs/2110.02378

index coding problem addresses constructions of broadcast functions that distribute information to the vertices of the graph whereby each vertex has access to the "side information " stored on the vertices in its neighborhood in the graph.

Essentially the same problem as that of storage coding, in a different guise, appears in a line of works devoted to guessing games on graphs, e.g., [9, 8], which has developed independently of both the storage codes and index coding problems. That these groups of problems are largely equivalent was realized in a number of papers, and the historical development is detailed in [2] from the index codes' perspective.

The main question concerning storage codes is determining the largest possible cardinality of the code $\mathcal{C}$ for a given graph G. Below we denote by $R_q(G)$ the maximum possible rate $\frac{1}{N}\log_q|\mathcal{C}|$ of a q-ary storage code for a given graph G. Determining $R_q(G)$ (or even $R(G) := \sup_q R_q(G)$ if we wish to optimize on the choice of the code alphabet) is a difficult problem related to the so-called minrank of the graph, and we refer to [16] for the best known bounds as well as an overview of the earlier results. Below we limit ourselves to finite field alphabets $Q = \mathbb{F}_q$ and assume that the code $\mathcal{C}$ forms a *linear subspace* of $\mathbb{F}_q^N$.

Let us start with the remark that if G is the complete graph, then a storage code on G is simply a 1-erasure correcting code, for example the parity code with rate $R_q(G) = 1 - 1/N$. For general graphs let us mention three generic constructions:

Matching construction A matching $M \subset E$ in a graph $G(V, E)$ defines a linear storage code $\mathcal{C}\colon \mathbb{F}_q^{|M|} \to \mathbb{F}_q^N$ by associating to $(x_e)_{e\in M}$ the vector $c \in \mathbb{F}_q^N$ such that for every vertex v incident to an edge $e \in M$ we assign $c_v = x_e$ and for every remaining vertex v we put $c_v = 0$. In particular, if G is d-regular and bipartite, it contains a perfect matching, giving rise to a storage code $\mathcal{C}(G)$ of rate $1/2$ independently of q.

Edge-vertex construction If the graph G does not have a perfect matching, we may nevertheless obtain a storage code of rate $1/2$ on a d-regular graph $G(V, E)$ in the following way. Consider the space $(\mathbb{F}_q)^{|E|}$ of vectors indexed by the edges of G. Next, map this space on $(\mathbb{F}_q^d)^N \cong \mathbb{F}_{q^d}^N$ by assigning to each vertex v a d-vector formed of the symbols written on the edges incident to it. In other words, $\mathcal{C}(G) = \{(c_v, v \in V)\}$ is obtained as the image of the mapping

$$\begin{aligned} \mathbb{F}_q^{|E|} &\to \mathbb{F}_{q^d}^N \\ (x_e)_{e\in E} &\mapsto (c_v)_{v\in V}, \text{ where } c_v = (x_e)_{e\in\partial(v)}, \end{aligned} \tag{1}$$

where $\partial(v)$ is the edge neighborhood of v in G. Since $|(\mathbb{F}_q)^{|E|}| = q^{\frac{dN}{2}}$, the rate of $\mathcal{C}$ is indeed 1/2.

Clique-vertex construction The edge-vertex construction affords a straightforward generalization if every vertex $v \in V$ is incident to the same number (say m) of k-cliques, where $k > 2$. Let $\mathcal{K}$ be the set of k-cliques in G and let us map $(\mathbb{F}_q^{k-1})^{|\mathcal{K}|}$ to $(\mathbb{F}_q^m)^n$ by placing a q-ary code of length k and dimension $k-1$ on every clique and distributing the symbols of the k-codeword to the vertices that form the clique. The rate of the code $\mathcal{C}(G)$ obtained as a result equals $(k-1)/k$; for instance, a triangular lattice gives rise to a code of rate $2/3$, etc.

For an overview and details of known constructions, see [9, 8, 16] and also [3, Ch.6].

Let us just give one upper bound on the rate of a storage code. It is known that $R_q(G)$ satisfies the constraints

$$M(G) \leq NR_q(G) \leq N - \alpha(G) \tag{2}$$

where $M(G)$ is the size of the largest matching in G and $\alpha(G)$ is the independence number of G. This result was proved in [15] for storage codes and in an earlier independent work [9] in the language of guessing games.

Note that if the graph G is a regular bipartite graph, then the upper and lower bound in (2) coincide, showing that one cannot improve upon the simple matching construction, and that for this class of graphs the maximum rate is $1/2$. We note also that if the graph G is *triangle-free* (i.e. has no cliques of size ≥ 3) then the above generic constructions (and others not mentioned here) all yield storage codes of rate at most $1/2$. This motivates the question of the existence of storage codes of rate $> 1/2$ for triangle-free graphs. This last point was made in particular in [5] (in the language of index codes) and [8] (in the language of guessing games on graphs).

2 Storage codes on triangle-free graphs

Both [5] and [8] reported only one example[2] of a storage code on triangle-free graphs for which $R_2(G) > \frac{1}{2}$. This code is associated to the graph Γ shown below in Fig. 1 (the authors of [8] also gave several examples of triangle-free graphs with $R_q(G) > \frac{1}{2}$ for $q > 2$). This graph can be defined in several ways and is known as the *Clebsch graph* or a *folded cube*, see Fig. 1. The storage code constructed on Γ is a linear binary code $\mathcal{C}$ of length $N = 16$ whose parity-check matrix has the form

$$\tilde{A} = I_N + A(\Gamma), \tag{3}$$

where $A(\Gamma)$ is the adjacency matrix of Γ. The identity matrix I_N is added since A includes only the neighborhood $\mathcal{N}(v)$ but not the vertex v itself. Upon checking that the $\mathbb{F}_2$-rank of $\tilde{A}$ equals 6, we conclude that the dimension of the code equals $N - \operatorname{rk}\tilde{A} = 10$, so $R_2(G) \geq \frac{5}{8}$. This example refuted a conjecture in [9] which suggested that the fractional clique cover bound holds with equality for triangle-free graphs. It is also easy to check that Γ contains independent sets of size 5, and thus from (2) we conclude that $\frac{5}{8} \leq R_2(\Gamma) \leq \frac{11}{16}$. Note that the recovery functions of the vertices use full neighborhoods (the parity-check equations have weight 6), even though the general definition of the storage code does not include this constraint. We call binary codes from this subclass *full-parity* storage codes, and it is only such codes that we consider from now on.

The Clebsch graph forms a rather special example: it is a unique strongly regular graph with the parameters $(16, 5, 0, 2)$; see [11], Theorem 10.6.4. Six other triangle-free strongly regular

[2]If we have a graph G that admits a storage code C of rate $R_q(G)$, we can construct many more graphs with the same storage rate, simply by taking k disjoint copies of the original graph G. The associated storage code will consist of the cartesian product of k copies of the original code C, whose codewords consist therefore of k successive codewords of C. We can also add arbitrary edges to and between the copies of G that the storage code will simply ignore. We will however conflate a code with its successive cartesian powers, considering that we are dealing with the same code ‘up to repetitions’.

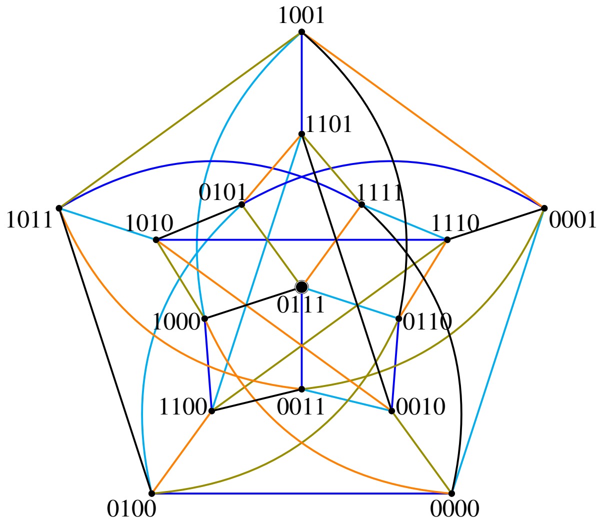

Figure 1: The Clebsch graph Γ (a folded cube): A 5-regular triangle-free graph on $n = 16$ vertices obtained by identifying each pair of opposite vertices in the 5-dimensional hypercube. It is also a coset graph of the binary code $\{00000, 11111\}$.

graphs are known [6, p. 119], and all the other examples of (nonbinary) storage codes of rate greater than $\frac{1}{2}$ in [8] are drawn from this list.

2.1 Storage codes on Cayley graphs

Recall that the Cayley graph $\mathsf{Cay}(\mathcal{G}, S)$ of the group $\mathcal{G}$ for a given set of generators S has $\mathcal{G}$ as its set of vertices, and the vertices g_1 and g_2 are connected by an edge if there is an element $s \in S$ such that $g_1 s = g_2$. For the group $\mathcal{G} = \mathbb{F}_2^r$, which is the only group we will consider, any subset S coincides with its inverse S^{-1}, so the graphs that we study are undirected. Since the generators are r-dimensional binary vectors, we also write the group action additively. Write the elements of S as columns of an $r \times n$ matrix H and consider the binary code C defined by H as the parity-check matrix. The graph $\mathsf{Cay}(\mathcal{G}, S)$ can be also viewed as a *coset graph* of the code C constructed on the set of cosets in $\mathbb{F}_2^n/C$, wherein two cosets are connected with an edge if and only if the Hamming distance between them (as subsets) is one.

Observe that the Clebsch graph Γ can be viewed as a Cayley graph over the group $\mathbb{F}_2^4$. Namely, consider the set S whose elements form the matrix

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{4}$$

Indeed, as easily checked, the canonical generators e_1, e_2, e_3, e_4 connect a vertex $x \in \mathbb{F}_2^4$ to its neighbors at Hamming distance one, while their sum $e_1 + \cdots + e_4$ connects it to its opposite vertex. See Fig. 1 for one possible vertex labeling, where each color corresponds to the action of a specific generator. Viewing the matrix H as a parity-check matrix of a $[5, 1]$ binary repetition code, we see that Γ is in fact the *coset graph* of this code.

Motivated by this example, we now investigate other Cayley graphs over binary groups $\mathbb{F}_2^r$ and exhibit more exceptional graph families that admit binary storage codes of rate greater than $1/2$. These graphs have connections to both classical and quantum coding theory.

Notation

Our notation is as follows. We are given a binary code C with parameters $[n, k, d]$ with a fixed parity-check matrix H which we are free to choose. Next we construct the coset graph $G = \mathsf{Cay}(\mathbb{F}_2^r, S)$ where $r = n - k$ is the number of rows and S is the set of columns of H. Note that G is a regular graph of degree $|S| = n$. The adjacency matrix A of this graph is symmetric, of order $N = 2^r$, and its rows and columns are labeled by the vectors $x \in \mathbb{F}_2^r$: we would like row x to specify the parity-check equation that recovers the value of the vertex x from its neighbors. Note however that $A_{x,x} = 0$, therefore, to involve the value supported by x in the parity-check equation, we consider, as in (3), the matrix

$$\tilde{A} = I_N + A(G). \tag{5}$$

Our storage code is finally a linear space in $\mathbb{F}_2^N$ defined as $\mathcal{C} = \ker(\tilde{A})$[3]. Note that the matrix $\tilde{A}$ is the adjacency matrix of the graph $\mathsf{Cay}(\mathbb{F}_2^r, S)$ to which we have added self-loops at every vertex, or equivalently the Cayley graph $\mathsf{Cay}(\mathbb{F}_2^r, S')$, where S' is obtained from S by adding 0 to the set of generators. The main problem addressed below is analyzing the dimension of $\mathcal{C}$ both in general and for several specific constructions.

As our first observation, note that once the minimum distance of the code C is at least 4, then so is the girth of the graph $\mathsf{Cay}(\mathbb{F}_2^r, S)$, i.e., the graph is triangle-free. We will therefore assume that all the small codes below have distance 4 or more.

The next lemma will help to further motivate our problem.

Lemma 1. *Let A be the adjacency matrix of the graph $\mathsf{Cay}(\mathbb{F}_2^r, S)$ where S may or may not contain 0. If $n = |S|$ is odd then $\operatorname{rk} A = N = 2^r$ and if n is even then $AA^\intercal = 0$, implying in particular $\operatorname{rk} A \le N/2$.*

Proof. The rows and columns of A are indexed by vectors of $\mathbb{F}_2^r$. Two distinct rows x and y intersect in the set of positions $\mathcal{I} = (x + S) \cap (y + S)$ which is of even size because whenever $x + s_1 = y + s_2$ is in $\mathcal{I}$, so is $x + s_2 = y + s_1$, implying that $\mathcal{I}$ is partitioned into pairs of the form $\{z, z + (x + y)\}$. Therefore the matrix $AA^\intercal$ has only zeros outside the main diagonal. Now a diagonal element has value $\langle x, x\rangle = |S| \bmod 2$ so when n is odd we have $AA^\intercal = I_N$ and when n is even we have $AA^\intercal = 0$, hence the claims of the lemma. $\square$

As a consequence, we observe that for a set of non-zero generators of *odd size* n, we have $\operatorname{rk} \tilde{A} \le N/2$, whence $\dim\ker(\tilde{A}) \ge N/2$ so the rate of the storage code satisfies $R(\mathcal{C}) \ge 1/2$. It may happen that for some Cayley graphs of odd degree (not counting the loops) we have

[3] Notice that we deal with two types of binary codes: the codes in $\mathbb{F}_2^n$ and codes in $\mathbb{F}_2^N$, called small codes and big codes in [10].

$\operatorname{rk}\tilde{A} < N/2$ in which case we will obtain an exceptional storage code of large rate. While this observation shows that this construction has a potential for uncovering large-size storage codes, it does not necessarily make finding such codes a straightforward task: in fact, as we have mentioned, for some years it was believed that they did not exist at all.

2.2 Coset graphs of binary codes: the quantum coding connection

To introduce this connection, recall that any binary linear code $\mathcal{C} \subset \mathbb{F}_2^N$ generated by a matrix A such that $AA^\intercal = 0$ (i.e. $\mathcal{C}$ is a self-orthogonal code) defines a quantum code through a particular case of the CSS construction [7, 18]. The quantum code has parameters $[[N, N - 2\operatorname{rk} A, D]]$, i.e. length N, dimension $N - 2\operatorname{rk} A$ and minimum distance D equal to the smallest Hamming weight of a vector in $\mathcal{C}^\perp \setminus \mathcal{C}$.

Therefore, for Cayley graphs over $\mathbb{F}_2^r$, the storage codes we are interested in also define quantum codes. From the point of view of storage, we do not have much use for the quantum code's minimum distance, but we are very much interested in its dimension: in particular we will obtain a storage code of rate greater than $1/2$ if and only if the associated quantum code has non-zero dimension. Quantum codes arising from Cayley graphs over $\mathbb{F}_2^r$ were studied in [10] and we will make use of some of results obtained in that work.

We now focus on the case of coset graphs of repetition codes of odd length $n = r + 1$, since they form a natural generalization of the Clebsch graph which is the coset graph of the $[5, 1]$ repetition code. Alternatively, they can be seen as Cayley graphs $\mathsf{Cay}(\mathbb{F}_2^r, S)$, r even, with generating set $S = \{e_1, e_2, \dots, e_r, e_1 + e_2 + \cdots + e_r\}$, where $e_1, \dots, e_r$ denote the canonical generators. Now it turns out that in a quantum coding context, coset graphs G_m of repetition codes of *even* length m were studied in [10], because their adjacency matrices $A(G_m)$ are self-orthogonal and therefore directly yield quantum codes. These quantum codes were proved in [10] to have parameters $[[2^{m-1}, 2^{m/2}, 2^{m/2-1}]]$. We use this result to prove:

Proposition 2. *The storage codes associated to coset graphs of repetition codes of odd length $n \geq 5$ have length $N = 2^{n-1}$ and dimension $K = 2^{n-2} + 2^{(n-3)/2}$.*

This last proposition therefore yields an infinite family of graphs for which the associated storage codes have rate exceeding $1/2$. In particular, for $n = 5$ we recover the value of the rate $\frac{5}{8}$. As n increases, the rate

$$R(G_m) = \frac{1}{2} + \frac{1}{2^{(n+1)/2}}$$

decreases to 1/2, so the highest rate within this family is achieved for $n = 5$, i.e. for the Clebsch graph.

2.3 Necessary conditions for high rate of codes on coset graphs

The starting observation in this part is given by the following proposition, due to Lowzow [13], as is point (a) of Theorem 4 below, that was formulated in a quantum coding context.

Proposition 3. *Let V be a vector subspace of $\mathbb{F}_2^r$. If $|S \cap V|$ is odd, then $\operatorname{rk} A \geq 2^{\dim V}$.*

Definition 1. *The* Schur product *of two codes $A, B \subset \mathbb{F}_2^n$ is a binary linear code $C = A * B \subset \mathbb{F}_2^n$ generated by all coordinatewise products $a * b = (a_1 b_1, \dots, a_n b_n), a = (a_1, \dots, a_n) \in A, b = (b_1, \dots, b_n) \in B$.*

A consequence of Proposition 3 is the following theorem:

Theorem 4. *Let $G = \mathsf{Cay}(\mathbb{F}_2^r, S)$ be a Cayley graph over $\mathbb{F}_2^r$ with set of generators S containing the 0 element. Let H be the $r \times n$ matrix whose columns are made up of the non-zero elements of S, so that $n = |S| - 1$, and let C be a code of length n with parity-check matrix H. Finally, let $\mathcal{C}$ be the storage code on G. Then*

(a) If $R(\mathcal{C}) > 1/2$, then n is odd and all the rows of H have even weight;

(b) If $R(\mathcal{C}) > (2^k - 1)/2^k, k = 2, 3, \dots$, then $(C^\perp)^{*(k-1)} \subset C$.

2.4 A family of storage codes on triangle-free graphs of rate $3/4$

We consider a family of storage codes on coset graphs of a specially constructed family of binary codes C_r of length $n = 2^{r-1} + 1$, dimension $k = 2^{r-1} - r$, and distance 4. To define it, start with the parity-check matrix of the extended Hamming code of length 2^{r-1}, augment it with an all-zero column, and then add a row of weight 2 that contains a '1' in the last position. Denote the resulting $(r+1) \times n$ matrix by H_r. For instance, for $r = 4$ we obtain the matrix

$$H_4 = \left[\begin{array}{cccccccc|c} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \tag{6}$$

where the choice of the last row is largely arbitrary as long as it is of weight 2 and intersects the all-zero column.

Observe that the code C_r does not contain its dual $C_r^\perp$, i.e., the code generated by H_r, because for instance the last row of H_r is not orthogonal to the other rows. Thus, from Theorem 4, the most we can hope for the storage code constructed on the coset graph of C_r is rate $3/4$. Our main result consists of proving that the rate $R(\mathcal{C}_r)$ is in fact close to this maximum value.

Theorem 5. *Let $\mathcal{C}_r$ be the $[N = 2^{r+1}, K]$ storage code constructed on the coset graph of the code $C_r, r \geq 4$. Then*

$$\frac{K}{N} = \frac{3}{4} - \frac{1}{2^r}.$$

The idea behind the construction is that the kernel of the adjacency operator associated to the coset graph of the extended Hamming code is easily described, and that the matrix (6) is essentially obtained by adding an extra row to the parity-check matrix of the extended Hamming code: we can therefore hope to compute the dimension of the kernel of the adjacency operator associated to this new code from the adjacency operator for the coset graph of the extended Hamming code. Now the latter graph is simply a complete bipartite graph whose adjacency matrix is

$$A = \begin{bmatrix} 0 & J \\ J & 0 \end{bmatrix}$$

where J is the all-one matrix of order 2^{r-1}. The kernel of A is formed of vectors of length 2^r of even weight in the first half as well as in the second half, and therefore has dimension $2^r - 2$. To switch from the coset graph of the extended Hamming code to the coset graph of the code C_r we need to understand the effect on the adjacency operators of adding an extra coordinate to a set of generators. For this we derived a formula that helps to derive the new dimension of the kernel of the adjacency operator when an extra coordinate is added to the set of generators. For details see [4].

2.5 Numerical experiments

It is tempting to look for other coset graphs of linear codes whose full-parity storage code has rate greater than $1/2$. It is a challenge to derive general formulae for dimensions however. With computer help we find:

Proposition 6.

(a) The binary Golay code of length 23 *and dimension* 11 *yields a graph* G *on* $N = 2048$ *vertices that supports a linear storage code of rate* $41/64$.

(b) The 2-error-correcting binary BCH code of length $n = 2^s - 1$ *and dimension* $k = 2^s - 1 - 2s$ *yields a graph* G_s *with* $N = 2^{2s}$ *vertices that supports a linear storage code with rate given in the following table*

s	4	5	6	7	8
$R_2(G_s)$	$\frac{39}{64}$	$\frac{347}{512}$	$\frac{1497}{2048}$	$\frac{6387}{8192}$	$\frac{26859}{32768}$.

The sequence of rate values obtained in Proposition 6 is $0.6094, 0.6777, 0.7309, 0.7796, 0.8196$, and the value $R_2(G_8) = 0.8196$ of the storage code of length $N = 65536$ now represents the largest known rate of storage codes on triangle-free graphs over any alphabet.

2.6 Open problems

1. *The rate question:* Arguably, the central question is whether rates of storage codes on triangle-free graphs can be arbitrarily close to 1. If not, what is an upper limit for these rates ?

2. *BCH codes:* In view of the numerical experiments it is of interest to find or estimate the dimension of storage codes obtained from the 2-error-correcting BCH codes. More specifically, what is $\limsup_{s\to\infty} R_2(G_s)$? At this point we cannot even rule out that the sequence $(R_2(G_s))_s$ in the limit reaches 1, which would resolve the question of the maximum possible rate of storage codes on triangle-free graphs.

3. *Reed-Muller codes:* Another good candidate is a subcode of the second order Reed-Muller code. Namely, for a given $m \geq 3$ consider the set of functions $v_i, 1 \leq i \leq m$ and $v_i v_j, 1 \leq i < j \leq m$. Evaluating these functions on the nonzero points of $\mathbb{F}_2^m$, we obtain a set of vectors of length $2^m - 1$. The linear code C_m spanned by these vectors forms a subcode of the punctured Reed-Muller code $RM(m, 2)$. Another way of constructing this code is to take a linear space formed of the Simplex code S_m and its Schur square $S_m * S_m$. The distance of the code C is 2^{m-2}, so its coset graph is triangle-free.

This choice of the subcode gives odd length and even weights of the rows, and the code C_m also contains its dual, as well as Schur powers of its dual, fulfilling the necessary conditions for high-rate storage codes. Finding the actual rate is however not straightforward, and we leave this as an open problem.

ACKNOWLEDGMENT

The research of Alexander Barg was partially supported by NSF grants CCF2110113 and CCF2104489.

References

[1] N. Alon, E. Lubetzky, U. Stav, A. Weinstein, and A. Hassidim. Broadcasting with side information. In *Proc. 49th IEEE Symp. Found. Comput. Science*, pages 823–832, 2008.

[2] F. Arbabjolfaei and Y. Kim. Three stories on a two-sided coin: Index coding, locally recoverable distributed storage, and guessing games on graphs. In *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 843–850, 2015.

[3] F. Arbabjolfaei and Y.-H. Kim. Fundamentals of index coding. *Foundations and Trends in Comm. Inf. Theory*, 14(3-4):163–346, 2018.

[4] A. Barg and G. Zémor. Full version of this paper. available at `https://arxiv.org/abs/2110.02378`.

[5] A. Blasiak, R. Kleinberg, and E. Lubetzky. Broadcasting with side information: Bounding and approximating the broadcast rate. *IEEE Trans. Inform. Theory*, 59(9):5811–5823, 2013.

[6] A.E. Brouwer and W.H. Haemers. *Spectra of Graphs*. Springer, New York e.a., 2012.

[7] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54:1098–1105, 1996.

[8] P. J. Cameron, A. N. Dang, and S. Riis. Guessing games on triangle-free graphs. *Electron. J. Combin.*, 23(1):Paper 48, 2016.

[9] D. Christofides and K. Markström. The guessing number of undirected graphs. *Electron. J. Combin.*, 18(1):Paper 192, 2011.

[10] A. Couvreur, N. Delfosse, and G. Zémor. A construction of quantum LDPC codes from Cayley graphs. *IEEE Trans. Inform. Theory*, 59(9):6087–6098, 2013.

[11] C. Godsil and G. Royle. *Algebraic Graph Theory*. Springer-Verlag, New York e.a., 2001.

[12] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin. On the locality of codeword symbols. *IEEE Trans. Inform. Theory*, 58(11):6925–6934, 2011.

[13] C. L. Lowzow. On quantum CSS-codes from Cayley graphs. Master's thesis, Université de Bordeaux, 2020.

[14] A. Mazumdar. On a duality between recoverable distributed storage and index coding. In *Proc. 2014 IEEE Int. Sympos. Inf. Theory*, pages 1977–1981, 2014.

[15] A. Mazumdar. Storage capacity of repairable networks. *IEEE Transactions on Information Theory*, 61(11):5810–5821, 2015.

[16] A. Mazumdar, A. McGregor, and S. Vorotnikova. Storage capacity as an information-theoretic vertex cover and the index coding rate. *IEEE Trans. Inform. Theory*, 65(9):5580–5591, 2019.

[17] K. Shanmugam and A. G. Dimakis. Bounding multiple unicasts through index coding and locally repairable codes. In *2014 IEEE International Symposium on Information Theory*, pages 296–300, 2014.

[18] A. Steane. Multiple-particle interference and quantum error correction. *Proc. R. Soc. Lond. A*, 452:2551–2577, 1996.