Designs, Codes and Cryptography (2022) 90:939–945
https://doi.org/10.1007/s10623-022-01020-8

A construction of maximally recoverable codes

Alexander Barg[1,2] · **Zitan Chen**[3] · **Itzhak Tamo**[4]

Received: 18 September 2021 / Revised: 13 January 2022 / Accepted: 4 February 2022 /
Published online: 16 February 2022

Abstract
We construct a family of linear maximally recoverable codes with locality r and dimension $r+1$. For codes of length n with $r \approx n^{\alpha}$, $0 \le \alpha \le 1$ the code alphabet is of the order $n^{1+3\alpha}$, which improves upon the previously known constructions of maximally recoverable codes.

Keywords Distributed storage · Codes with local recovery · Maximally recoverable codes

Mathematics Subject Classification 94B60

1 Introduction

Consider a linear code $\mathcal{C}$ over a finite field $F = \mathbb{F}_q$ of length n and dimension k, and let r be a number such that $r+1$ divides n. We will write $[n] = \{(i, j), j = 1, \dots, r+1;\ i = 1, \dots, \frac{n}{r+1}\}$, and for $i = 1, \dots, \frac{n}{r+1}$ we will call the subset of indices $R_i = \{(i, j), j = 1, \dots, r+1\}$ a *repair group*. Call a set $T \subset [n]$ a transversal of the set of repair groups $\mathscr{R} = (R_i)_i$ if $|T \cap R_i| = 1$ for all i. For a subset $X \subset [n]$ denote by $\mathcal{C}|_X$ the puncturing of $\mathcal{C}$ in the coordinates in X, i.e., a coordinate projection of $\mathcal{C}$ on the complementary subset $X^c := [n] \backslash X$.

Communicated by V. A. Zinoviev.

✉ Alexander Barg
abarg@umd.edu

Zitan Chen
chenztan@gmail.com

Itzhak Tamo
zactamo@gmail.com

1 Department of ECE and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA

2 Inst. for Probl. Inform. Trans., Moscow, Russia

3 School of Science and Engineering and Future Network of Intelligence Institute, The Chinese University of Hong Kong, Shenzhen, China

4 Department of EE-Systems, Tel Aviv University, Tel Aviv, Israel

Definition 1 The code $\mathcal{C}$ is called *maximally recoverable* (MR) with locality r if the following two properties hold:

(i) For any repair group R_i the code $\mathcal{C}|_{R_i^c}$ has distance at least 2;
(ii) For any transversal T of $\mathscr{R}$ the code $\mathcal{C}|_T$ is maximum distance separable.

We write the parameters of the code $\mathcal{C}$ as $(n, k, r) =$ (length, dimension, locality).

This definition is a particular case of a more general notion of MR codes introduced in [5]. Namely, one assumes that (i), any repair group is capable of correcting any $a \geq 1$ erasures, and (ii), upon puncturing any a coordinates from each of the repair groups, the resulting code is a maximum distance separable code that can correct h erasures. Thus, Definition 1 corresponds to the case of $a = 1$ and $h = \frac{nr}{r+1} - k$.

The main problems related to MR codes are: the minimum field size q required to construct an MR code with a given set of parameters, and finding explicit constructions of MR codes, with sizeable literature devoted to them over the last decade. In this note we prove the following result.[1]

Theorem 1 *There exists a family of* $(n, k = r + 1, r)$ *MR codes over* $\mathbb{F}_q$, *where* $r = \Theta(n^\alpha)$, $0 \leq \alpha \leq 1$, *with* $q = \Theta(n^{1+3\alpha+o(1)})$.

To establish this theorem, we develop an idea behind one construction in [6] that gave codes with locality $r = 3$ and large k relying on Behrend's classic result on sets of integers with no 3-term arithmetic progressions [2].

For the context we include a sample of the known results on the construction and parameters of MR codes, focusing on the regime of large h relevant to us. Among the known families of MR codes we note constructions with $q = O((r+1)n^{2h-1})$ [4], $q = O((r+1)^{\frac{nr}{r+1}})$ [9], $q = O(\max(r+1, \frac{n}{r+1}))^h$ [3], as well as a number of constructions in [7] with comparable parameters. We refer the reader to the introduction of [6, 7], or [3], Table 1, for a more detailed overview which also covers the entire range of possible parameters a, r, and h. As remarked in [7], most of the known constructions require alphabets of size q that depend exponentially on h. One exception where the minimum alphabet size is independent of h is the construction of [8] which requires $q = O(\max(n/(r+1), r+1))^r$. This is larger than the result in Theorem 1 above both for fixed and growing r. In summary, the code family constructed in this paper improves upon the known results in terms of the required field size.

To address the question of lower bounds (impossibility), we observe that the known nontrivial results [6] assume that h is a fixed constant. At the same time, in our setting h is clearly increasing with n, and the only known constraints on q are general bounds of the form $q = \Omega(n)$ (for instance, part (ii) of Definition 1 implies that $q \geq k+1$, see [5, Theorem 19]).

2 The construction

Let γ be a primitive element of F and let $N = q - 1$ be the size of its multiplicative group. We will define a linear code over F by specifying its generator matrix $G = (g_{\ell,(i,j)})$ of dimensions $(r+1) \times n$ where the length n will be determined later. Let

[1] We use standard asymptotic notation: for functions $f(n), g(n), n \in \mathbb{N}$ we write $f(n) = O(g(n))$ if $f(n) \leq Cg(n)$ for some constant C starting with some n; $f(n) = \Omega(g(n))$ if $f(n) \geq cg(n)$ starting with some n, and $f(n) = \Theta(g(n))$ if both $f(n) = O(g(n))$ and $g(n) = O(f(n))$.

$$\mathcal{A} = \bigcup_{i=1}^{\frac{n}{r+1}} A_i \subset \mathbb{Z}_N$$

be a subset of size n formed as a union of pairwise disjoint sets A_i, and let $\{a_{ij}, j = 1, \ldots, r+1\}$ be the elements of A_i, $i = 1, \ldots, n/(r+1)$. Define

$$g_{\ell,(i,j)} = \begin{cases} \gamma^{\ell \cdot a_{ij}} & 1 \le \ell \le r \\ \gamma^{(r+1)\cdot a_{ij}} + (-1)^{r+1} & \ell = r+1. \end{cases} \tag{1}$$

The main property of the set $\mathcal{A}$ that supports the construction is the following: let $\{a_1, \ldots, a_{r+1}\} \subset \mathcal{A}$, then

$$\sum_{s=1}^{r+1} a_s = 0 \pmod N \text{ if and only if exists } i \in \{1, \ldots, \tfrac{n}{r+1}\} \text{ s.t. } A_i = \{a_1, \ldots, a_{r+1}\}. \tag{2}$$

Theorem 2 *Let G be the matrix defined above, where the set $\mathcal{A}$ satisfies (2). Then the rows of G span an $(n, r+1, r)$-MR code over F.*

Next we give a construction of the set $\mathcal{A}$ with the required properties. Let λ and δ satisfy $0 < \lambda < \frac{1}{r^3}, 0 < \delta < \frac{\lambda}{r}$, and define $d = \lfloor \delta N \rfloor, l = \lfloor \lambda N \rfloor$. Suppose that $D \subset \{1, \ldots, d\}$ is a subset of integers such that for any $d_0, \ldots, d_r \in D$ the equation over $\mathbb{Z}$

$$d_0 + \cdots + d_{r-1} = r d_r \tag{3}$$

is satisfied only if $d_0 = d_1 = \cdots = d_r$. Define $r+1$ subsets $D_i \subset \mathbb{Z}_N$ by letting

$$D_i = \begin{cases} il + D & 0 \le i \le r-1 \\ N - \binom{r}{2} l - r \cdot D & i = r, \end{cases} \tag{4}$$

where $b + D, b \cdot D$ mean adding or multiplying every element of D by b. By the choice of λ and δ one can verify that the subsets D_i are disjoint. Define the set $\mathcal{A} = \cup_{i=0}^{r} D_i$ and note that $|\mathcal{A}| = n := |D|(r+1)$. Consider a partition of $\mathcal{A}$ into disjoint transversals A_b for any $b \in D$, where

$$A_b = \{a_{i,b} : i = 0, \ldots, r\} \text{ and } a_{i,b} = \begin{cases} il + b, & 0 \le i \le r-1 \\ N - \binom{r}{2} l - rb, & i = r. \end{cases} \tag{5}$$

Lemma 1 *The partition $\mathcal{A} = \cup_{b \in D} A_b$ satisfies property (2).*

Large sets of integers that satisfy (3) exist, namely, the following is true.

Lemma 2 [1, Lemma 3.1] *For every $r \ge 2$ and every positive integer m, there exists a subset $D \subset \{1, 2, \ldots, m\}$ of size at least*

$$|D| \ge \frac{m}{e^{5\sqrt{\log m \log r}}}$$

that has property (3).

This claim is proved by an averaging argument over intersections of a subset of integers with spheres of varying radii, and this is the only non-explicit part of our construction. We include a short proof at the end of the next section to make the presentation self-contained.

Putting things together, we have constructed an $(n, r+1, r)$ MR code $\mathcal{C}$ of length

$$n = |D|(r+1) \geq (r+1)d\, e^{-5\sqrt{\log d \log r}},$$

where we have used Lemma 2 with $m = d$. Let us estimate the dependence of the field size q on the parameters of $\mathcal{C}$, letting $n, q \to \infty$. We have

$$\log n \geq \Omega\Big(\log q - 3\log r - 5\sqrt{\log \frac{q}{r^4} \log r}\Big). \tag{6}$$

where we put $d = \Theta(\frac{q}{r^4})$ (this appears to be the best choice given our assumption on δ). Suppose that $r = \Theta(n^\alpha)$, where $0 \leq \alpha \leq 1$ and note that this includes the cases of constant r and various rates of increase of r up to $r = \Theta(n)$, i.e., a constant number of repair groups. Now from (6) we obtain the estimate for q stated in Theorem 1.

3 Proofs

Proof of Theorem 2 Let $S \subseteq [n/(r+1)] \times [r+1]$ be an $(r+1)$-subset of indices and let G_S be a square submatrix of G of order $r+1$ whose columns are indexed by the elements of S (since S is a set of pairs, this definition is consistent with (1)). We begin by showing that the rank of G_S is r if S forms a repair group, otherwise, G_S is of full rank. First note that the first r rows of G_S form an $r \times (r+1)$ Vandermonde submatrix, hence the rank of G_S is at least r. The rank is exactly r if and only if there exists a nonzero vector $f = (f_1, \ldots, f_r, f_{r+1})$ such that $f \cdot G_S = 0$. Note that $f_{r+1} \neq 0$ since otherwise, it would violate the fact that the first r rows of G_S are linearly independent. Therefore, assume wlog that $f_{r+1} = 1$. Since the columns of G are defined by elements γ^β, $\beta \in \mathcal{A}$, the conditions $f \cdot G_S = 0$ are alternatively written as $f(\gamma^{\beta_i}) = 0$ for $i = 1, \cdots, r+1$ and some $\beta_i \in \mathcal{A}$, where

$$f(x) := x^{r+1} + (-1)^{r+1} + \sum_{i=1}^{r} f_i x^i.$$

By assumption, the monic polynomial $f(x)$ has $r+1$ zeros γ^{β_i}, and thus

$$f(x) = \prod_{i=1}^{r+1} (x - \gamma^{\beta_i}).$$

By comparing the constant terms in the two expressions of f we have

$$(-1)^{r+1} = \prod_{i=1}^{r+1} (-\gamma^{\beta_i}) = (-1)^{r+1} \gamma^{\sum_{i=1}^{r+1} \beta_i}$$

or $\sum_{i=1}^{r+1} \beta_i = 0 (\mathrm{mod}\, N)$. Then, by recalling assumption (2), either the subset S forms a repair group, or otherwise G_S is of full rank. Hence property (ii) in Definition 1 holds. Next, assume that S forms a repair group, and we need to show that $\mathcal{C}|_{S^c}$ has distance at least 2. To prove this, note that any $(r+1) \times r$ submatrix of G_S has rank r since it contains an $r \times r$ Vandermonde submatrix. Since $\mathrm{rk}(G_S) = r$, any column of G_S is in the span of the remaining r columns, and thus the code $\mathcal{C}|_{S^c}$ corrects a single erasure. □

Proof of Lemma 1 Let $\mathscr{B} := \{b_0, b_1, \ldots, b_r\} \subset \mathcal{A}$. We will show that

$$\sum_{i=0}^{r} b_i = 0 \ (\mathrm{mod}\, N). \tag{7}$$

is satisfied if and only if $\mathscr{B}$ coincides with one of the transversals A_b defined in (5). One direction is easy: namely, the elements in every A_b sum to 0 modulo N. Indeed

$$\sum_{j=0}^{r} a_{i,b} = \sum_{j=0}^{r-1}(il+b) + N - \binom{r}{2}l - rb = 0 \pmod{N}.$$

Conversely, suppose (7) holds. We aim to prove that $\mathscr{B} = A_b$ for some $b \in D$. Let $t = |\mathscr{B} \cap D_r|$ and let us first show that $t = 1$. Indeed, if $t = 0$, then each $b_i \notin D_r$ and therefore $b_i \le (r-1)l + d$ over $\mathbb{Z}$, and (again over $\mathbb{Z}$)

$$0 < \sum_{i=0}^{r} b_i \le r((r-1)l+d) < r^2 l < N,$$

where the last inequality follows by the choice of λ. This contradicts (7), so $t \ge 1$. A similar argument applies in the case of $t \ge 2$, namely we will show that in such a case $(t-1)N < \sum_{i=0}^{r} b_i < tN$ over $\mathbb{Z}$, which again will contradict (7). Indeed, we have

$$\begin{aligned}\sum_{i=0}^{r} b_i &> \sum_{b_i \in D_r} b_i \ge tN - t\binom{r}{2}l - trd \ge (t-1)N + N\Big(1 - t\lambda\Big(\binom{r}{2}+1\Big)\Big)\\ &= (t-1)N + N\Big(1 - t\lambda\frac{r^2-r+2}{2}\Big) > (t-1)N,\end{aligned}$$

where the last step follows since $\lambda < r^{-3}$, $t \le r+1$ and $r \ge 2$. For the upper bound write

$$\begin{aligned}\sum_{i=0}^{r} b_i &= \sum_{b_i \in D_r} b_i + \sum_{b_i \notin D_r} b_i\\ &\le tN - t\binom{r}{2}l + (r+1-t)((r-1)l+d)\\ &\le tN - 2\binom{r}{2}l + (r-1)((r-1)l+d)\\ &= tN - (r-1)(l-d) < tN,\end{aligned}$$

again contradicting (7). We conclude that $t = 1$ and suppose that $D_r \cap \mathscr{B} = \{b_r\}$, where

$$b_r = N - \binom{r}{2}l - rz, \tag{8}$$

for some $z \in D$.

Our next goal is to show that all the other elements in $\mathscr{B}$ are of the form $b_i = ie + z, i = 0, \ldots, r-1$, and here property (3) comes in handy. We begin with arguing that $t_i := |\mathscr{B} \cap D_i| = 1$ for all $i = 0, \ldots, r-1$. Note that over $\mathbb{Z}$

$$\sum_{i=0}^{r-1} b_i \le r((r-1)l+d) < r^2 l < N,$$

hence from (7) and (8)

$$\sum_{i=0}^{r-1} b_i = \binom{r}{2}l + rz,$$

again over $\mathbb{Z}$. Clearly, $\sum_{i=0}^{r-1} t_i = r$, and we will show that

$$\sum_{i=0}^{r-1} i t_i = \binom{r}{2}. \tag{9}$$

Indeed, if $\sum_{i=0}^{r-1} i t_i \geq \binom{r}{2} + 1$, then

$$\begin{aligned} l\left(\binom{r}{2}+1\right) &> l\binom{r}{2} + rd \geq \binom{r}{2} l + rz = \sum_{i=0}^{r-1} b_i \\ &= \sum_{i=0}^{r-1} \sum_{b \in \mathscr{B} \cap D_i} b \geq \sum_{i=0}^{r-1} \sum_{b \in \mathscr{B} \cap D_i} il = \sum_{i=0}^{r-1} i t_i l \geq l\left(\binom{r}{2}+1\right), \end{aligned} \tag{10}$$

and we arrive at a contradiction. Similarly, if $\sum_{i=0}^{r-1} i t_i \leq \binom{r}{2} - 1$, then

$$\begin{aligned} \binom{r}{2} l &< \binom{r}{2} l + rz = \sum_{i=0}^{r-1} \sum_{b \in \mathscr{B} \cap D_i} b \leq \sum_{i=0}^{r-1} \sum_{b \in \mathscr{B} \cap D_i} (il + d) \\ &= \sum_{i=0}^{r-1} t_i (il + d) \leq \left(\binom{r}{2} - 1\right) l + rd < \binom{r}{2} l, \end{aligned} \tag{11}$$

and (11) makes no sense, and thus (9) holds.

Finally, recalling (4), let

$$\mathscr{B} \cap D_i = \{il + b_{i,j}, 1 \leq j \leq t_i\}$$

where the $b_{i,j}$'s are t_i distinct elements of D. Then over $\mathbb{Z}$

$$\binom{r}{2} l + rz = \sum_{i=0}^{r-1} b_i = \sum_{i=0}^{r-1} \sum_{j=1}^{t_i} (il + b_{i,j}) = \binom{r}{2} l + \sum_{i=0}^{r-1} \sum_{j=1}^{t_i} b_{i,j},$$

hence

$$\sum_{i=0}^{r-1} \sum_{j=1}^{t_i} b_{i,j} = rz.$$

Now (3) implies that $b_{i,j} = z$ for all i, j. However, the numbers b_i were chosen distinct, and thus, $t_i = 1$ for all $i \leq r$. Moreover, $b_i = il + z$, $i = 0, 1, \ldots, r-1$. On account of (8) and (5) the proof is complete. □

Proof of Lemma 2 We closely follow [1], adding some details. Let h be an integer, to be chosen later. Consider a set of integer numbers $D = (x_i)_i$ written in the form $x_i = \sum_{j=0}^{t} x_{i,j} h^j$, where $0 \leq x_{i,j} < \frac{h}{r}$, $i = 0, 1, \ldots, t$, $t = \lfloor \log_h m \rfloor - 1$, and suppose further that for every $x_i \in D$

$$\sum_{j=0}^{t} x_{i,j}^2 = B.$$

If an $(r+1)$-tuple $x_0, x_1, \ldots, x_{r+1}$ satisfies (3), then for every $j = 0, 1, \ldots, t$

$$x_{0,j} + x_{1,j} + \cdots + x_{r-1,j} = r x_{r,j}. \tag{12}$$

By the convexity of the function $z \mapsto z^2$ this implies that

$$x_{0,j}^2 + x_{1,j}^2 + \cdots + x_{r-1,j}^2 \geq r x_{r,j}^2,$$

with equality if and only if $x_{0,j} = x_{1,j} = \cdots = x_{r,j}$. At the same time,

$$\sum_{i=0}^{r-1} \sum_{j=0}^{t} x_{i,j}^2 = rB = r \sum_{j=0}^{t} x_{r,j}^2.$$

The last two relations imply that only identical $(r+1)$-tuples satisfy (12), and thus only identical $(r+1)$-tuples of elements in D satisfy (3).

Clearly, $B \leq (t+1)\frac{h^2}{r^2}$, so there is a choice of B such that

$$|D| \geq \frac{h^{t+1}}{r^{t+1}(t+1)\frac{h^2}{r^2}} \geq \frac{m}{h^3 r^{t-1}(t+1)}.$$

Take $h = \lfloor e^{\sqrt{\log m \log r}} \rfloor$, then $(t-1)\log r < \sqrt{\log m \log r}$ and

$$h^3 r^{t-1}(t+1) \leq e^{5\sqrt{\log m \log r}}.$$

□

Acknowledgements Alexander Barg was partially supported by NSF-BSF grant CCF2110113 and NSF grant CCF2104489. Itzhak Tamo was supported by the European Research Council (ERC Grant No. 852953) and by the Israel Science Foundation (ISF Grant No. 1030/15).

References

1. Alon N.: Testing subgraphs in large graphs. In: Proceedings 42nd IEEE Symposium on Foundations of Computer Science, pp. 434–441, (2001).
2. Behrend F.A.: On sets of integers which contain no three terms in arithmetical progression. Proc. Nat. Acad. Sci. USA **32**, 331–332 (1946).
3. Cai H., Miao Y., Schwartz M., Tang X.: A construction of maximally recoverable codes with order-optimal field size. IEEE Trans. Inf. Theory **68**(1), 204–212 (2022).
4. Gabrys R., Yaakobi E., Blaum M., Siegel P.H.: Constructions of partial MDS codes over small fields. IEEE Trans. Inf. Theory **65**(6), 3692–3701 (2019).
5. Gopalan P., Huang C., Jenkins B., Yekhanin S.: Explicit maximally recoverable codes with locality. IEEE Trans. Inf. Theory **60**(9), 5245–5256 (2014).
6. Gopi S., Guruswami V., Yekhanin S.: Maximally recoverable LRCs: a field size lower bound and constructions for few heavy parities. IEEE Trans. Inf. Theory **66**(10), 6066–6083 (2020).
7. Guruswami V., Jin L., Xing C.: Constructions of maximally recoverable local reconstruction codes via function fields. IEEE Trans. Inf. Theory **66**(10), 6133–6143 (2020).
8. Martínez-Peñas U., Kschischang F.R.: Universal and dynamic locally repairable codes with maximal recoverability via sum-rank codes. IEEE Trans. Inf. Theory **65**(12), 7790–7805 (2019).
9. Neri A., Horlemann-Trautmann A.-L.: Random construction of partial MDS codes. Des. Codes Cryptogr. **88**(4), 711–725 (2020).