Characterizations and construction methods for linear functional-repair storage codes

Junming Ke

University of Tartu

Algebra Seminar, March 2022

Table of Contents

Table of Contents

Linear storage codes

A *linear* exact-repair storage code is a collection of n subspaces $U_1, ..., U_n$ of an $m-$dimentional vector space $\mathbb{F}^m$, each of dimension α.

- $\mathbb{F}^m$: message space
- U_i: (storage) node spaces
- α: (storage) node capacity

Linear storage codes

A *Recovery set* of the storage code is a subset of the storage node spaces that together span the entire message space $\mathbb{F}^m$.
The *span* of a collection of vector spaces $W_1, ..., W_k$ is the collection of all vectors $w_1 + ... + w_k$ with $w_i \in W_i, \forall i$ (that is, the smallest vector space containing all the vector spaces $W_1, ..., W_k$.).

- *Recovery dimension*: the smallest k of the recovery set.

Linear storage codes

Using linear storage codes.

- Associate the n storage nodes $v_1, ..., v_n$ with the n storage spaces $U_1, ..., U_n$, and choose a fixed basis in each of them.
- Represent the data to be stored as an vector $x \in \mathbb{F}^m$.
- In node i, we store the α inner products of x with the α basis vectors of U_i.

Linear storage codes

If B_i is the $m \times \alpha$ matrix that has as its columns the basis vectors for U_i, then the storage node v_i stores $x^{\perp} B_i$.

- The DSS can compute the inner product of the data vector x and any vector contained in one of the node spaces;
- From the data stored in a recovery set, which by definition is a spanning subset of the node spaces, the DSS can recover x.

Repair problem

A collection R of the node spaces is a *repair set* for a certain node space $U_l \notin R$ if it is possible to choose in each $U_i \in R$ a $\beta-$dimensional *repair space* $W_{i,l}$ such that U_l is contained in the span of the repair space $W_{i,l}$.

- β: fix some positive integer as *transport capacity*.
- The vector $x^{\perp}B_l$ stored in node v_l can be recovered from a linear combination of the vectors $x^{\perp}B_{i,l}$.

Repair problem

If *each* node space in the code has the repair set of size r with respect to transport capacity β then we say that the storage code has *repair locality* r with respect to transport capacity β.

- We refer to the above storage code as a linear exact-repair $(m; n, k, r, \alpha, \beta)$-*storage code*.
- The coding rate $R = \frac{m}{n\alpha}$.

An example of linear storage codes

Examples

$U_0 =< e_0, e_2 + e_3 >$,
$U_1 =< e_1, e_3 + e_0 >$,
$U_2 =< e_2, e_0 + e_1 >$,
$U_3 =< e_3, e_1 + e_2 >$,
considered as subspaces of $\mathbb{F}_2^4$.
These node spaces constitute an
$(m = 4; n = 4, k = 2, r = 3, \alpha = 2, \beta = 1)$ linear exact-repair storage code.

Remark

$R = \{U_1, U_2, U_3\}$ is a repair set for node space U_0.
The linear transformation given by $e_i \to e_{i+1}$ (indices modulo 4) maps U_i to U_{i+1}. (symmetry.)

Table of Contents

Linear functional-repair storage codes

For linear exact-repair storage codes, a *fixed* vector space was associated with each node.
Linear *functional-repair* storage codes can be thought as a specification of *admissable* node space arrangements, with the property that in every such arrangement, a node space can be "repaired" by replacing it with a (possibly different) space so that the resulting arrangement again satisfies the specifications.

An example of linear functional-repair storage codes

Examples

Consider subspaces of $\mathbb{F}_2^5$. We will construct an $(m = 5; n = 4, k = r = 3, \alpha = 2, \beta = 1)$ linear functional-repair storage code.

At any moment, the four 2-*dimensional* node spaces $U_1, ..., U_4$ associated with the four storage nodes satisfy the following specification:

1. Any two of the node spaces intersect trivially, $U_i \cap U_j = \{0\}$ when $i \neq j$.
2. Any three of the node spaces span the entire message space $\mathbb{F}^5$.

An example of linear functional-repair storage codes

Examples

$U_1 =< e_1, a_1 >$,
$U_2 =< e_2, a_2 >$,
$U_3 =< e_3, a_3 >$,
for some basis e_1, e_2, e_3, a_1, a_2 of $\mathbb{F}^5$, with $a_1 + a_2 + a_3 = 0$.
Suppose that node 4 fails.
We can take $U_4 =< e_1 + e_2, e_1 + e_3 >$.

Remark

Given $\beta = 1$ we must choose a vector $w_i \in U_i^* = U_i \backslash \{0\}$ for $i = 1, 2, 3$ and let U_4 be some 2-dimensional subspace (Rule 1) of their span w_1, w_2, w_3.
$U_4 = \{0, w_1 + w_2, w_1 + w_3, w_2 + w_3\}$, $w_3 \neq a_1 + a_2, w_1 \neq a_1, w_2 \neq a_2$ (Rule 2).
So $w_1 = e_1 + x_1 a_1, w_2 = e_2 + x_2 a_2, w_3 = e_3 + x_3 a_3$, $x_1, x_2, x_3 \in \mathbb{F}_2$.

Table of Contents

Admissible states

Consider the previous example.
The collection $\mathcal{U} = \{U_1, U_2, U_3\}$ satisfies the specification and that a fourth node space U_4 can be constructed such that any triple from $\mathcal{U} \cup \{U_4\}$ again satisfies the specification (admissible repairing collections).

Admissible states

Definition 3.1

Let $\mathcal{U}$ be a collection of α-dimensional subspaces of a vector space $\mathbb{F}^m$. We say that an α-dimensional subspace U of $\mathbb{F}^m$ can be obtained from $\mathcal{U}$ by (r, β)-*repair* if it is possible to choose r spaces $U_1, ..., U_r$ in $\mathcal{U}$, the *repair set*, and then a β-dimensional subspace $W_i \subseteq U_i$ in each of them, the *repair spaces*, such that U is contained in the span $W_1 + ... + W_r$ of the repair spaces.

Admissible states

Definition 3.2

A linear functional repair storage code with parameters $(m; n, k, r, \alpha, \beta)$ over some finite field $\mathbb{F}$ is a set $\mathcal{A}$ of $(n-1)$-tuples $\mathcal{U}$ of α-dimensional subspaces of $\mathbb{F}^m$, such that the following *repair property* holds.
Given any $(n-1)$-tuple $\mathcal{U} = U_1, ..., U_{n-1}$ in $\mathcal{A}$, there exists an α-dimensional space U that can be obtained from $\mathcal{U}$ by (r, β)-repair such that for every $i = 1, ..., n-1$, the $(n-1)$-tuple $\mathcal{U} \cup \{U\} \backslash \{U_i\}$ is again in $\mathcal{A}$. Furthermore, we require that each $(n-1)$-tuple $\mathcal{U}$ in $\mathcal{A}$ contains a *spanning subset* of size k, that is, there can be found k spaces in $\mathcal{U}$ that together span $\mathbb{F}^m$.

Repair problem

- *admissible repairing collections* of the code: $(n-1)$-tuples $\mathcal{U}$ in $\mathcal{A}$.
- *admissible states* of the code: $(\mathcal{U}, U)$.

Examples

The first example with admissible states $(\{U_1, ..., U_n\} \backslash \{U_i\}, U_i)$.
The second example (same admissible states) with as admissible repairing collections all triples $\mathcal{U} = \{U_1, U_2, U_3\}$ of 2-dimensional node spaces in $\mathbb{F}^5$ that together span $\mathbb{F}^5$ in which any pair of node spaces intersecting trivially.

Admissible states

Theorem 3.3

Every linear functional-repair code according to Definition 3.2 is indeed a storage code under the regime of functional repair. Conversely, every linear functional-repair code can be obtained from a collection $\mathcal{A}$ of admissible repairing collections with the properties as in Definition 3.2.

Table of Contents

Family of functional-repair codes

A collection of $(s+1)$-dimensional vector spaces $\mathcal{U} = \{U_1, ..., U_r\}$ in $\mathbb{F}^m$ with $m = m_{r,s}$ is (r, s)-*good* if the span of any $r - s + j$ of these subspaces has dimension $(r-s)(s+1) + s + ... + (s+1-j)$, for every $j = 0, ..., s$.
We will prove the existence of a family of codes with parameters $(m_{r,s}; n = r+1, k = r, r, \alpha = s+1, \beta = 1)$, where $m = m_{r,s} = (r-s)(s+1) + s + ... + 1 = (r-s)(s+1) + \binom{s+1}{2}$ and $r > s \geq 0$.

Examples

The second example is the case where $r = 3, s = 1$.

Family of functional-repair codes

Theorem 4.1

Suppose that $\mathcal{U} = \{U_1, ..., U_r\}$ is (r, s)-good over $\mathbb{F}$. Let $w_i \in U_i$ for $i = 1, ..., r$ and let U be an α-dimensional subspace of the span W of $w_1, ..., w_r$. Let $C \subseteq \mathbb{F}^r$ be the collection of all vectors $c = (c_1, ..., c_r)$ for which $\sum_{j=1}^{r} c_j w_j \in U$. Then the collections $\mathcal{U} \cup \{U\} \backslash \{U_i\}$ for all $i = 1, ..., r$ are all (r, s)-good if and only if the vectors $w_1, ..., w_r$ are independent and the code C is an $[r, s+1, r-s]$ linear Maximum Distance Separable (MDS) code over $\mathbb{F}$.

Family of functional-repair codes

It has been shown that the mentioned MDS codes certainly exists for a field size $|\mathbb{F}| \geq r - 1$ [1].
It follows immediately from the (r, s)-good property that independent vectors $w_1, ..., w_r$ as in the theorem can always be found, moreover, they can be used to recursively construct at least one (r, s)-good collection for all r and s with $r > s \geq 0$.
Theorem 4.1 and theorem 3.3 prove the existence of linear functional-repair codes for the above parameters, with field size equal to smallest prime power q for which $q \geq r - 1$.
[1] MacWilliams, Florence Jessie, and Neil James Alexander Sloane. The theory of error correcting codes. Vol. 16. Elsevier, 1977.

Table of Contents

Construction method using groups

The set of admissible repairing collections that define the code and the number of admissible states can be huge, which severely complicates the data management in the DSS. Therefore, it is desirable to find *small* codes. For example, if possible we would like to find a small *subset* of admissible repairing collections that itself again defines a code, or we would like to find a small set of admissible states for certain given parameters *directly*.

Construction method using groups

Theorem 5.1

Let $\sigma = (\mathcal{U} = \{U_1, ..., U_{n-1}, U = U_n\})$, with each of $U_1, ..., U_n$ an α-dimensional subspace of $\mathbb{F}^m$, such that U_n can be obtained by (r, β)-repair from $\mathcal{U}$, and with $\mathcal{U}$ containing a spanning k-subset if $k < n - 1$. Suppose we can find for every $i = 1, ..., n - 1$ an invertible linear map L_i mapping $\mathcal{U}$ to $\mathcal{U} \cup \{U_n\} \backslash \{U_i\}$. Then with $\mathcal{G} =< L_1, ..., L_n >$, the group generated by $L_1, ..., L_n$, the set A consisting of all images $G(\mathcal{U})$ with $G \in \mathcal{G}$ is a collection of admissible repairing collections for a linear storage code with parameters $(m; n, k, r, \alpha, \beta)$.

Table of Contents

Combinatorial description

A *vector space partition* for a vector space V is a collection of subspaces $V_1, ..., V_m$, not necessarily all of the same dimension, such that each nonzero vector in V is contained in exactly one of the spaces V_i. It turns out that our code arises from a vector space partition of Beutelspacher type [2].

[2] Cullina, Daniel, Alexandros G. Dimakis, and Tracey Ho. "Searching for minimum storage regenerating codes." arXiv preprint arXiv:0910.2245 (2009).

Combinatorial description

Consider the finite field $W = \mathbb{F}_8$, constructed with a primitive element α with $\alpha^3 = \alpha + 1$.
Consider W as a 3-dimensional vector space over $\mathbb{F}_2$. Note that W has a 2-dimensional subspace $U = \{0, \alpha, \alpha^2, \alpha^4\}$ of W that is invariant under the Frobenius map $x \to x^2$.
We will construct a vector space partition for $V = W \oplus U$.
For each $\beta \in \mathbb{F}_8$, we define $U_\beta = \{(\beta u, u) | u \in U\}$.
There are 31 nonzero vectors in V, 7 of which are in W and another 3 in each of the 8 subspaces U_β.

Combinatorial description

See the original paper. [3]

[3] Hollmann, Henk DL, and Wencin Poh. "Characterizations and construction methods for linear functional-repair storage codes." 2013 IEEE International Symposium on Information Theory. IEEE, 2013.

Table of Contents

Subsequential work

This work derives the trade-off between the storage capacity and the repair bandwidth for partial repair of multiple failed nodes (based on information flow graph). The authors present explicit codes that achieve the optimal cut-set bound for functional repair of multiple nodes with high probability, the proposed scheme employs a subpacketization level that scales linearly in the code parameters. [4]

[4] Mital, Nitish, et al. "Functional Broadcast Repair of Multiple Partial Failures in Wireless Distributed Storage Systems." IEEE Journal on Selected Areas in Information Theory 2.4 (2021): 1093-1107.

Subsequential work

This work applies novel coding techniques to edge storage and computation problems, and shows similar significant gains in terms of storage costs, communication costs, computation costs, and privacy. [5]

[5] Mital, Nitish. "Coding-theoretic approaches to distributed caching, storage and computing." (2020).

Subsequential work

This work presents a code construction and repair scheme for optimal functional regeneration of multiple node failures, which is based on stitching together short MDS codes on carefully chosen sets of points lying on a linearized polynomial. [6]

[6] Mital, Nitish, et al. "Practical functional regenerating codes for broadcast repair of multiple nodes." 2019 IEEE International Symposium on Information Theory (ISIT). IEEE, 2019.

Subsequential work

This work shows that the existing predominant coding theoretic and vector space models of repair codes can be given a unified treatment in a projective geometric framework, which permits a natural treatment of results such as the cutset bound. [7]

[7] Ng, Siaw-Lynn, and Maura B. Paterson. "Functional repair codes: a view from projective geometry." Designs, Codes and Cryptography 87.11 (2019): 2701-2722.

Subsequential work

This work studies the maximum virtual server rate of a PIR array code with the k-PIR property. the authors presents upper bounds on the achievable virtual server rate, some constructions, and ideas how to obtain the PIR array codes with the highest possible virtual server rate. [8]

[8] Blackburn, Simon R., and Tuvi Etzion. "PIR array codes with optimal virtual server rate." IEEE Transactions on Information Theory 65.10 (2019): 6136-6145.

Subsequential work

This work proposes a new simple optimal repair strategy for $(k + m, k)$ Hadamard MSR codes, which can considerably reduce the computation compared with the original one during the node repair. [9]

[9] Tang, Xiaohu, et al. "A new repair strategy for the Hadamard minimum storage regenerating codes for distributed storage systems." IEEE Transactions on Information Theory 61.10 (2015): 5271-5279.

Subsequential work

This work presents some new information theoretical lower bounds on the storage overhead as a function of the repair locality, valid for most common coding and repair models. [10]

[10] Hollmann, Henk DL. "On the minimum storage overhead of distributed storage codes with a given repair locality." 2014 IEEE International Symposium on Information Theory. IEEE, 2014.