ON APPLICATIONS OF ORBIT CODES TO STORAGE

SHIQIU LIU AND FRÉDÉRIQUE OGGIER
Division of Mathematical Sciences
School of Physical and Mathematical Sciences
Nanyang Technological University, Singapore

(Communicated by Paolo Vettori)

Abstract. We consider orbit codes, namely subspace codes obtained from orbits of the action of subgroups of $GL_n(F_q)$ on $m$-dimensional subspaces of $F_q^n$. We discuss their applicability to design storage codes, in particular in the context of collaborative repair: we identify known storage codes that can be interpreted as orbit codes, motivate why orbit codes provide good candidates for storage codes, and translate the storage code parameters into those of the algebraic objects involved. Two simple families of storage orbit codes are given.

1. Introduction

Consider an $(N, n)$ linear code $C$ over the finite field $F_q$, $q$ a prime power, that is, $C$ is an $n$-dimensional subspace of $F_q^N$. A codeword in $C$, in the context of communication, is sent over an erasure channel, which will result in an erroneous codeword, where some of its components have been erased. The code should be designed so as to permit the recovery of the codeword, assuming a given erasure pattern. The minimum Hamming distance $d_H$ of the code characterizes the erasure recovery capacity of the code, sometimes then denoted as an $(N, n, d_H)$ code.

This scenario can be easily translated into the context of distributed storage systems. Instead of sending a codeword over an erasure channel, store the $N$ coefficients of a codeword of $C$ into $N$ storage nodes. Node failures become erasures, and whether the stored encoded object can be retrieved given a failure pattern depends on $d_H$.

There is however a fundamental difference between erasure codes for communication, and erasure codes for distributed storage: in the former case, once the altered codeword is received, it is not deteriorated anymore, while in the latter case, a storage system is meant to last over months or years, and failures continue to happen. Therefore, no matter how good a linear code is, it cannot protect data from failures, without maintenance [10]. Maintenance means that in case of node failures, it is necessary to repair the data that became unavailable, namely: codeword coefficients which are erased need to be recomputed, and stored at live nodes. This can be done using a decoding algorithm, but often, it is preferred to have more efficient ways to compute codeword coefficients, typically as a linear combination of a small number of other codeword coefficients. When several nodes jointly repair $t$ failures, $t \geq 2$, 

2010 Mathematics Subject Classification: Primary 20H20; Secondary 11T71.

Key words and phrases: Group action, orbit codes, partial repair, subspace codes, storage codes.

The research of S. Liu was supported by the Singapore National Research Foundation under Research Grant NRF-RF2009-07. The research of F. Oggier was supported by the MoE Tier-2 grant “eCODE: Erasure Codes for Datacenter Environments”.
the repair process is said to be collaborative \[2, 13\]. The metrics of interest to decide the efficacy of a storage code with respect to maintenance are typically the amount of bandwidth needed per repaired node, which is a function of data stored per node, and the number of nodes contacted per repair. The latter is often in tension with the optimal trade-off between repair bandwidth and storage overhead.

To better handle the maintenance process, it is often useful to consider an \((Nm, n)\) linear code \(C\), with generator matrix \((v_1, \ldots, v_{Nm}) \in \text{Mat}_{n \times Nm}(\mathbb{F}_q)\), with columns \(v_i \in \mathbb{F}_q^n\), \(i = 1, \ldots, Nm\) (This is also an approach used in the context of array codes, see e.g. \[15\]). Then a data object \(o \in \mathbb{F}_q^n\) is encoded into a codeword \((ov_1, \ldots, ov_{Nm}) \in \mathbb{F}_q^{1 \times Nm}\), and the \(Nm\) codeword coefficients are distributed into \(N\) storage nodes, each node storing \(m\) coefficients. This helps the maintenance process in that it allows the content of a node to be repaired with lesser communication cost than if the node content is considered as a single symbol. Each storage node is assumed to be equipped with computational power, therefore, a node storing \(ov_{i_1}, \ldots, ov_{i_m}\) is able to compute \(\mathbb{F}_q\)-linear combinations \(\sum_{j=1}^{m} \alpha_i ov_{i_j}\). From this point of view, it makes sense to therefore model each storage node as storing an \(m\)-dimensional subspace of \(\mathbb{F}_q^n\) with \(\mathbb{F}_q\)-basis \(v_{i_1}, \ldots, v_{i_m}\) (we may assume without loss of generality that \(v_{i_1}, \ldots, v_{i_m}\) are linearly independent), as explicitly suggested in \[1\].

With the above viewpoint, the problem of designing storage codes then becomes that of finding \(Nm\)-dimensional subspaces of \(\mathbb{F}_q^n\), which naturally brings in the Grassmannian \(G_q(n, m)\), which is by definition the set of all subspaces of \(\mathbb{F}_q^n\) of dimension \(m\). The question is then how to pick subspaces appropriately for the purpose of storage. A first tentative to treat storage codes in this manner was presented in \[8\], using Grassmanian graphs, which yielded some interesting constructions. A systematic approach to design subspace codes for storage has been recently proposed in \[12\], which relies on Plücker embeddings. Subspace codes were in general studied from a group action point of view in \[14\], which introduced the concept of orbit codes. In this paper, motivated by the natural interpretation of storage codes in terms of subspaces and by their properties, we adopt the approach of orbit codes, and question the applicability of such codes to distributed storage, hoping to exploit group actions as a way to characterize good storage codes.

We are particularly interested in the collaborative case, where nodes involved in the repair of \(t\) failures, \(t \geq 2\), download data from live nodes, and exchange some of their obtained data among a subset of each others. When this subset is not proper and involves all \(t - 1\) other nodes, it is referred to as collaborative, while partial collaboration \[5\] emphasizes that data is exchanged among a proper subset of strictly less than \(t - 1\) nodes. One reason for choosing this setting is that the one failure case has been already well studied and many codes are available (see \[10\] and reference therein), while the scenarios of collaboration, and especially partial collaboration, have been less studied, and few constructions are known for partial collaboration \[6\]. We focus on two classical extreme regimes: one where the amount of repair bandwidth per repaired node is minimized, and one where the amount of storage per node is minimized.

The paper is organized as follows. In Section 2, we translate code parameters in terms of orbit codes, and give a motivating example. In Section 3, we focus on orbit codes generated by the companion matrix of a polynomial, and propose a simple family of codes, for partial collaborative repair, which minimizes the repair bandwidth. We further identify a known class of storage codes \[9\] as being one
instance of cyclic orbit codes. In Section 4, we present another construction of orbit codes for partial collaboration, aiming this time at minimizing the storage. Though this instance of orbit codes is simple, it does achieve the optimal performance in terms of trade-off between storage and repair bandwidth. In Section 5, we compare the two constructions with known bounds: the first construction is shown to achieve the minimum repair bandwidth, but trade the corresponding optimal storage with the minimal number of nodes contacted per repair. Note that as of now, optimal families for partial collaboration at the minimum repair bandwidth case are not known. The second construction is indeed offering the minimum storage together with its corresponding minimum repair bandwidth. The number of contacted nodes per repair is large. The presented codes have the further benefit of being easily generated, as orbit codes are likely to be in general. It turns out that though the framework of orbit codes seem fairly natural to address storage codes, and some known storage codes can be interpreted as orbit codes, the dependency of the code parameters not only on the cyclic group which is acting but also on the original subspace on which the cyclic group acts complicates the study of such codes, which therefore leaves many open questions.

2. Storage codes from orbit codes

As above, let \( \mathbb{F}_q \) be the finite field with \( q \) elements, with \( q \) a prime power, and let \( G_q(m, n) \) be the set of all subspaces of \( \mathbb{F}_q^n \) of dimension \( m \), called Grassmannian:

\[
G_q(m, n) = \{ \mathcal{U} \text{ subspace of } \mathbb{F}_q^n, \dim(\mathcal{U}) = m \}.
\]

We denote by \( GL_n(\mathbb{F}_q) \) the set of \( n \times n \) invertible matrices with coefficients in \( \mathbb{F}_q \). Multiplication by elements of \( GL_n(\mathbb{F}_q) \) defines a group action from the right on \( G_q(m, n) \) by

\[
G_q(m, n) \times GL_n(\mathbb{F}_q) \rightarrow G_q(m, n)
\]

\[
(\mathcal{U}, A) \mapsto \mathcal{U}A = \{ uA, u \in \mathcal{U} \}
\]

since any element of \( GL_n(\mathbb{F}_q) \) maps an \( m \)-dimensional subspace to an \( m \)-dimensional subspace. In fact, as pointed out in [14], since any two \( m \)-dimensional subspaces can be mapped onto each other by an element of \( GL_n(\mathbb{F}_q) \), \( GL_n(\mathbb{F}_q) \) acts transitively on \( G_q(m, n) \).

Fix \( \mathcal{U} \in G_q(m, n) \), let \( G \) be a subgroup of \( GL_n(\mathbb{F}_q) \), and let \( \mathcal{U}G = \{ \mathcal{U}g, g \in G \} \) be the orbit of \( \mathcal{U} \) under the right action of \( G \). We will refer to \( \mathcal{U}G \) as an orbit code, following the terminology of [14]. If furthermore \( G \) is cyclic, we call \( \mathcal{U}G \) a cyclic orbit code.

In order to represent an \( m \)-dimensional subspace \( \mathcal{U} \) of the vector space \( \mathbb{F}_q^n \), we fix a basis of \( \mathcal{U} \), and use an \( m \times n \) matrix \( U \) whose row space \( \{ vU, v \in \mathbb{F}_q^m \} \) is \( \mathcal{U} \).

Orbit codes form collections of \( m \)-dimensional subspaces of \( \mathbb{F}_q^n \) obtained by group action, which have been successfully used in the context of network coding [14]. This makes them potential candidates to provide codes for distributed storage. Let us see how the storage parameters are translated into parameters for an orbit code \( \mathcal{U}G \), assuming that an object \( o \) of length \( n \) is stored.

2.1. Code length. The number \( N \) of storage nodes is the length of the code, that is the cardinality \( |\mathcal{U}G| \) of the orbit \( \mathcal{U}G \), and it is well known that

\[
|\mathcal{U}G| = \frac{|G|}{|\text{Stab}_G(\mathcal{U})|}.
\]
where $\text{Stab}_G(\mathcal{U}) = \{g \in G, \mathcal{U}g = \mathcal{U}\}$ is a subgroup of $G$ called the stabilizer of $\mathcal{U}$. Alternatively

$$|\mathcal{U}G| = \frac{|G|}{|G \cap \text{Stab}_{GL_n(F_q)}(\mathcal{U})|}.$$ 

Indeed, if $g \in \text{Stab}_G(\mathcal{U})$, then $g \in G$ and $g$ is also an element of $GL_n(F_q)$ such that $\mathcal{U}g = \mathcal{U}$. For the reverse inclusion, similarly, if $g \in G$ and $\mathcal{U}g = \mathcal{U}$, then $g \in \text{Stab}_G(\mathcal{U})$.

The following two results give a way to compute the stabilizer of $\mathcal{U}$, and in turn the number of nodes over which the storage code is used.

**Lemma 1.** Let $\mathcal{U}_I$ be the $m$-dimensional subspace generated by the first $m$ unit vectors, that is, it has $U_I = (I_m, 0_{m \times (n-m)})$ as an $F_q$-basis. Then

$$\text{Stab}_{GL_n(F_q)}(\mathcal{U}_I) = \left\{ \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}, A \in GL_m(F_q), B \in F_q^{(n-m) \times m}, C \in GL_{n-m}(F_q) \right\}.$$ 

**Proof.** Take $g$ in $\text{Stab}_{GL_n(F_q)}(\mathcal{U}_I)$. Note that $U_Ig = (A, 0_{m \times (n-m)})$ and since $A$ is invertible, then $(A, 0)$ generates $\mathcal{U}_I$. Then $B$ can take any value, but $C$ needs to be invertible, to ensure that $g$ is. \qed 

We keep the above notation in what follows.

**Lemma 2.** For any $m$-dimensional subspace $\mathcal{U}$, we have

$$\text{Stab}_{GL_n(F_q)}(\mathcal{U}) = L^{-1}\text{Stab}_{GL_n(F_q)}(\mathcal{U}_I)L,$$

for $L$ such that $\mathcal{U} = U_IL$, $L \in GL_n(F_q)$.

**Proof.** Note first that any $m$-dimensional subspace $\mathcal{U}$ is of the form $\mathcal{U} = U_IL$ for some $L$ in $GL_n(F_q)$. Take $g \in \text{Stab}_{GL_n(F_q)}(\mathcal{U})$, then by definition

$$\mathcal{U}g = \mathcal{U} \iff U_ILg = U_IL \iff U_ILgL^{-1} = U_I,$$

showing that $gL^{-1} \in \text{Stab}_{GL_n(F_q)}(\mathcal{U}_I)$, that is $g \in L^{-1}\text{Stab}_{GL_n(F_q)}(\mathcal{U}_I)L$.

Conversely, take an element $L^{-1}gL \in L^{-1}\text{Stab}_{GL_n(F_q)}(\mathcal{U}_I)L$, then

$$\mathcal{U}L^{-1}gL = U_IL^{-1}gL = U_ILg = U_IL = \mathcal{U}$$

and $L^{-1}gL$ indeed belongs to the stabilizer of $\mathcal{U}$. \qed 

Using the above lemma, we have that

$$|\text{Stab}_{GL_n(F_q)}(\mathcal{U})| = |L^{-1}\text{Stab}_{GL_n(F_q)}(\mathcal{U}_I)L| = |\text{Stab}_{GL_n(F_q)}(\mathcal{U}_I)|.$$ 

2.2. Storage Capacity. The storage capacity (or number of stored symbols in $F_q$) for every node is $m$. This is because a subspace $\mathcal{U}$ of dimension $m$ means that the corresponding node actually stores $Uo^T \in F_q^m$, where we recall that $o \in F_q^{1 \times n}$ is the data object to be stored. This notion was introduced in [1]. Under the hypothesis that a stored object $o \in F_q^{1 \times n}$ should be recoverable from any $k$ nodes, the minimal value of $m$ is $n/k$. 

---

**Advances in Mathematics of Communications** Volume 10, No. 1 (2016), 113–130
2.3. Fault tolerance. The ability to recover a stored object \( o \in \mathbb{F}^1 \times n \) from any \( k \) nodes, thus providing protection against any \( N - k \) failures is rephrased as:

\[
\dim(U_{i_1} + \ldots + U_{i_k}) = n.
\]

Note that when \( m = 1 \), then, recalling the connection with classical codes presented in the introduction, every node stores one column of a generator matrix of a linear code, and the property of fault tolerance just discussed is that of a maximum distance separable (MDS) code.

For arbitrary \( k \), there is no formula to evaluate \( \dim(U_{i_1} + \ldots + U_{i_k}) \), apart from using iteratively the well known formula when \( k = 2 \) which states that

\[
(2) \quad \dim(U_{g_i} + U_{g_j}) = \dim(U_{g_i}) + \dim(U_{g_j}) - \dim(U_{g_i} \cap U_{g_j}).
\]

**Example 1.** Consider the Grassmannian \( G_2(3,5) \), and let \( G = \langle g \rangle \) be the subgroup of \( GL_5(\mathbb{F}_2) \) generated by

\[
g = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}.
\]

Consider the orbit code \( U_G \), where \( U = U_I \) has for basis

\[
U = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0
\end{bmatrix}.
\]

The length \( N \) of the orbit code \( U_G \) is

\[
N = |U_G| = \frac{|G|}{|Stab_G(U)|} = 4.
\]

Indeed, \( g \) has order 4

\[
g^2 = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}, \quad g^3 = \begin{bmatrix}
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 0
\end{bmatrix}, \quad g^4 = I_4,
\]

and \( Stab_G(U) \) is trivial, since \( G \) intersects trivially \( Stab_{GL_5(\mathbb{F}_2)}(U) \) which is formed of matrices whose first \( 3 \times 3 \) block is invertible (see Lemma 1). The orbit code \( U_G \) is

\[
U, \quad \begin{bmatrix}
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1
\end{bmatrix}, \quad \begin{bmatrix}
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0
\end{bmatrix}, \quad \begin{bmatrix}
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0
\end{bmatrix}.
\]

This corresponds to a storage code where the size of the object \( o \) is \( n = 5 \), every storage node stores \( m = 3 \) symbols, e.g., the first node corresponding to the subspace \( U \) stores \((1,0,0,0,0)o^T = o_1, (0,1,0,0,0)o^T = o_2, (0,0,1,0,0)o^T = o_3\) for \( o = (o_1, \ldots, o_5) \in \mathbb{F}_2^5 \), giving the following storage allocation:

- Node 1: \( o_1, o_2, o_3 \)  
- Node 2: \( o_3, o_5, \sum_i o_i \)  
- Node 3: \( \sum_i o_i, o_2, o_4 \)  
- Node 4: \( o_4, o_1, o_5 \).

The data object \( o \) may be retrieved out of any two nodes, since, from (2)

\[
\dim(U_{g_i} + U_{g_j}) = 2 \dim(U_{g_i}) - 1 = 5
\]
for all \( i \neq j \). Indeed, two distinct subspaces intersect in a subspace of dimension 1.

2.4. Repair. As pointed out in the introduction, the main difference between a storage code and an erasure code is that a storage code should be amenable to repair, namely, the code should be such that the data stored at one node can be computed from a (small) subset of other nodes, without (necessarily) having to decode the object first. Concretely, a repair of one failure is done as follows: a node downloads data from several, possibly all live nodes, and from the data gathered, recomputes the codeword coefficient lost because of the failure. The parameters of importance are typically the communication cost of the repair, called repair bandwidth, which is also related to the number of live nodes contacted.

- Decomposing the stored data per node into \( m \) pieces is beneficial to the communication cost: it allows a repair to download part of the stored data. The optimal repair bandwidth per failed node is \( m \), since at least the content of a failed node must be communicated for a repair to be successful.
- To speed up repairs, contacting few live nodes is beneficial [10]. The optimal number of nodes is thus 2 in the case of one failure (since less than 2 means that a repetition code is used), or 1 in the (partial) collaborative case.

One may consider either repair of one failure at a time, or (possibly) wait for several failures to occur before repairing \( t \) of them together. Joint repair of multiple failures is often done collaboratively [2, 13] as follows: a subset of nodes handle the repair, they each download some amount of information, perform the necessary computations on their data, but also further exchange data among each other, thus involving \( t - 1 \) nodes in the exchange process. If strictly less than \( t - 1 \) nodes are involved in the exchange, we speak of partial collaboration. Collaboration is known to be beneficial to the repair process [2, 13] in terms of reducing the repair bandwidth.

Example 2. We continue Example 1, where we have for \( \mathbf{O} \) of size 5

- Node 1: \( o_1, o_2, o_3 \), Node 3: \( \sum_i o_i, o_2, o_4 \)
- Node 2: \( o_3, o_5, \sum_i o_i \), Node 4: \( o_4, o_1, o_5 \).

In case of one node failure, the subspace \( \mathcal{O}g^i \) may be computed from the knowledge of a subspace of dimension 1 from each of the other \( \mathcal{O}g^j, j \neq i \), which follows from the fact that \( \dim(\mathcal{O}g^i \cap \mathcal{O}g^j) = 1 \) for all \( i \neq j \). For example, if Node 1 were to fail, it can be recomputed using \( o_1 \) from Node 4, \( o_2 \) from Node 3, and \( o_3 \) from Node 2. This code instance has been reported in [11], not in the language of orbit codes, but as a code whose repair bandwidth is minimal.

While the motivation example presented above is optimal for the case of one failure, most of our focus next will be on the collaborative case.

3. Instances of Cyclic Orbit Codes Using Companion Matrices

We next propose a family of cyclic orbit codes based on companion matrices and compute the parameters of the corresponding storage codes.
Definition 1. Let \( p(x) = \sum_{i=0}^{n-1} p_i x^i + x^n \) be a monic polynomial in \( \mathbb{F}_q[x] \). Then the companion matrix \( M_p \) of \( p(x) \) is defined by
\[
M_p = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
-p_0 & -p_1 & -p_2 & \cdots & -p_{n-1}
\end{bmatrix} \in \mathbb{F}_q^{n \times n}.
\]

We first discuss basic properties of codes, namely length and fault tolerance, before proposing code constructions.

3.1. Basic properties. The following is probably well known and is proven for the sake of completeness.

Lemma 3. Let \( G = \langle g \rangle \) be the subgroup of \( GL_n(\mathbb{F}_q) \) generated by the \( n \times n \) companion matrix
\[
g = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
-p_0 & -p_1 & -p_2 & \cdots & -p_{n-1}
\end{bmatrix}
\]
of a polynomial \( p(x) = \sum_{i=0}^{n-1} p_i x^i + x^n \in \mathbb{F}_q[x] \), \( p_0 \neq 0 \). Then the order of \( g \) is the smallest positive integer \( l \) such that \( p(x) \) divides \( x^l - 1 \).

Proof. It is well known that \( p(x) \) is the minimal polynomial of the companion matrix \( M_p \) defined in (3), namely \( p(x) \) is the monic polynomial of least degree such that \( p(M_p) = 0 \). Now since \( p_0 \neq 0 \), there exists a positive integer \( e \leq q^n - 1 \) such that \( p(x) \) divides \( x^e - 1 \) (see [4, Lemma 3.1]), so take \( l \) to be the smallest positive integer with the property that \( p(x)|x^l - 1 \). Consequently, \( M_p^l = I_n \) if and only if \( p(x)|x^l - 1 \).

Corollary 1 ([7]). Let \( G = \langle g \rangle \) be the subgroup generated by the companion matrix of the polynomial \( p(x) = \sum_{i=0}^{n-1} x^i \in \mathbb{F}_q[x] \). Then \( |G| = n + 1 \).

Proof. We have \( (x - 1) \sum_{i=0}^{n} x^i = x^{n+1} - 1 \).}

One advantage to deal with companion matrices is that we are sure that no matter the choice of \( U \), the resulting code will not be trivial (that is, containing \( U \) only).

Proposition 1. Let \( g \) be the companion matrix of \( p(x) \) as in (4), and let \( U \) be an \( m \)-dimensional subspace in \( G_q(m,n) \). Then the orbit code \( U \langle g \rangle \) is not trivial \( (|U \langle g \rangle| \geq 2) \).

Proof. Consider \( g \) as a linear map from \( \mathbb{F}_q^n \) to \( \mathbb{F}_q^n \), and \( \mathbb{F}_q^n \) as an \( \mathbb{F}_q[x] \)-module over the polynomial ring \( \mathbb{F}_q[x] \) [3], its invariant subspaces are given by the Smith normal form of \( xI_n - g \), a diagonal matrix whose diagonal coefficients are called elementary divisors, which we denoted by \( d_i \). They are computed as follows. Compute the \( i \times i \)
minors of
\[ xI_n - g = \begin{bmatrix} x & -1 & 0 & \ldots & 0 \\ 0 & x & -1 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -1 \\ p_0 & p_1 & p_2 & \ldots & x + p_{n-1} \end{bmatrix} \]
and let \( d_i \) denote the greatest common divisor of all the \( i \times i \) minors. Then \( d_1 = d_2 = \ldots = d_{n-1} = \pm 1 \), which shows that there is no invariant subspace of dimension \( m, 1 \leq m \leq n - 1 \).

While any choice of \( U \) does give a non-trivial orbit code, different choices of \( U \) may give different sizes of orbit codes, as illustrated next.

**Example 3.** Choose \( G = \langle g \rangle \) with
\[ g = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \]
which is the companion matrix of \( x^4 + x + 1 \in \mathbb{F}_2[x] \). Since \( x^4 + x + 1 \) is an irreducible polynomial, let \( \omega \) be a root, that is \( \omega^4 = \omega + 1 \), and \( \mathbb{F}_2[x]/(x^4 + x + 1) \simeq \mathbb{F}_2(\omega) \simeq \mathbb{F}_{2^4} \). Since \( \omega \) is a primitive element of \( \mathbb{F}_{2^4} \), the order of \( G \) is 15.

Consider \( U_I \in G_2(2, 4) \) with canonical basis \( U_I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \). By Lemma 1, the stabilizer of \( U_I \) under \( G \) is trivial, since \( G \) intersects trivially \( \text{Stab}_{GL_4(\mathbb{F}_2)}(U_I) \), which is formed of matrices whose first \( 2 \times 2 \) block is invertible. Therefore the length of the code \( U_I G \) is 15. The length of the code can more directly be seen to be 15 by invoking Lemma 4 below.

To illustrate the fact that a change in the choice of \( U \) may change the length of the corresponding orbit code, we look for powers of \( g \) whose set of of invariant subspaces is non-empty. Take \( g^5 \), given by
\[ g^5 = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}, \]
and as in Proposition 1, and keeping the same notation, compute the Smith normal form of \( xI_4 - g^5 \), to find that \( d_1 = d_2 = 1 \), while \( d_3 = x^2 + x + 1 \), and \( d_4 = x^4 + x^2 + 1 \). Therefore the invariant factors are
\[ d_1 = 1, \quad \frac{d_2}{d_1} = 1, \quad \frac{d_3}{d_2} = x^2 + x + 1, \quad \frac{d_4}{d_3} = x^2 + x + 1, \]
revealing the presence of invariant subspaces of dimension 2. Thus, let \( U \in G_2(2, 4) \) and \( U = U_I L \), where
\[ L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \]
so that a basis is given by
\[ U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}. \]
Since $\mathcal{U}g^5 = \mathcal{U}, \mathcal{U}g^{10} = (\mathcal{U}g^5)g^5 = \mathcal{U}$, and the stabilizer of $\mathcal{U}$ under $G$ has cardinality 3, given by $\text{Stab}_G(\mathcal{U}) = \{I_4, g^5, g^{10}\}$, so that the orbit code $\mathcal{U}G$, with $\mathcal{U} = U I_L$, is of length 5, while $\mathcal{U}I_G$ is of length 15.

The following result tells the length of the considered orbit codes from companion matrices when $q = 2$.

**Lemma 4.** Let $g$ be the companion matrix of $p(x)$, with $g$ and $p(x)$ as in (4), $G = \langle g \rangle$, and let $\mathcal{U}_I \in G_q(m, n)$ be the subspace generated by a canonical basis as above. If $q = 2$, then the stabilizer $\text{Stab}_G(\mathcal{U}_I)$ is trivial. Consequently the size of the orbit code is the order of $g$.

**Proof.** Let $l$ be the order of $g$, and suppose that the stabilizer of $\mathcal{U}_I$ under $G = \langle g \rangle$ is not trivial, that is, there exists an $s$, $0 < s < l$, such that $\mathcal{U}_I g^s = \mathcal{U}_I$. Then by Lemma 1, there exist $A \in GL_m(F_q), B \in F_q^{(n-m)\times m}, C \in GL_{n-m}(F_q)$ such that

$$
\begin{bmatrix}
A \\
B \\
C
\end{bmatrix} = \begin{bmatrix}
a_2 & 0 \\
a_3 & 0 \\
\vdots & \vdots \\
a_{m} & 0 \\
b_1 & c_1
\end{bmatrix},
$$

where $a_i, b_i, c_i$ denote the respective $i$th row of $A, B$ and $C$. Equating coefficientwise shows that

$$
A = a_{11} I_m, \quad b_1 = 0, \quad c_1 = (a_{11}, 0, \ldots, 0).
$$

Repeating the same computation by noting that the first $m$ rows of $g^s \cdot g$ and $g \cdot g^s$ are equal, and given respectively by

$$
\begin{bmatrix}
0_{m \times 1} & A & 0_{m \times (n-m-1)}
\end{bmatrix},
$$

we get

$$
b_2 = 0, \quad c_2 = (0, a_{11}, 0, \ldots, 0),
$$

and more generally, comparing the first $m$ rows of $g^{s+t} \cdot g$ and $g \cdot g^{s+t}$, $t = 1, \ldots, n-m-1$, yields

$$
b_j = 0, \quad c_j = (0, \ldots, a_{11}, 0), \quad j = 1, \ldots, n-m,
$$

where only the $j$th coordinate of $c_j$ is non-zero.

Therefore

$$
g^s = a_{11} I_n,
$$

and since $q = 2$, the only choice is $a_{11} = 1$. This is a contradiction, since $g$ has order $l > s$. \[\square\]

Note that the same argument breaks at $g^s = a_{11} I_n$ for $q \neq 2$. For example, when $q = 3$, $p(x) = x^3 + 2$, then $g$ has order 6, and $g^3 = 2I_3$. 

- Advances in Mathematics of Communications Volume 10, No. 1 (2016), 113–130
3.2. Code construction I: $t$ failures with minimized bandwidth. We saw above that under the same group action, $UG$ and $U_2G$ may yield different code lengths, due to different stabilizers. In this section, we thus consider the canonical case $U_2G$.

**Lemma 5.** Take $U$, $U' \in G_q(n-1, n)$, $U \neq U'$. Then $U$ and $U'$ intersect in a subspace of dimension $n-2$.

**Proof.** Since by (2)
\[
\dim(U + U') = 2(n-1) - \dim(U \cap U') \leq n,
\]
it must be that
\[
n - 2 \leq \dim(U \cap U')
\]
which concludes the proof, since $U$ and $U'$ are distinct. \qed

The following lemma describes the behavior of intersection of three subspaces, and shows that the considered orbit codes exhibit a particular structure, namely any three subspaces intersect in a subspace of dimension $n - 3$.

**Lemma 6.** Take $U$, $U'$, $U''$ three distinct elements of $G_q(n-1, n)$. Then $U, U'$ and $U''$ intersect in a subspace of dimension either $n - 2$ or $n - 3$. When $U, U', U''$ belong to an orbit code $U(g)$ with $g$ the companion matrix of $p(x)$ as in (4), the dimension of their intersection is $n - 3$.

**Proof.** Since by (2)
\[
\dim(U \cap U' \cap U'') = \dim(U) + \dim(U' \cap U'') - \dim(U + (U' \cap U'')),
\]
and we know that $\dim(U) = n - 1$ and $\dim(U' \cap U'') = n - 2$ by the above lemma, we have that
\[
\dim(U \cap U' \cap U'') = 2n - 3 - \dim(U + (U' \cap U'')).
\]
Now either $(U' \cap U'') \subset U$ and $\dim(U + (U' \cap U'')) = n - 1$, or $(U' \cap U'') \not\subset U$ and $\dim(U + (U' \cap U'')) = n$, which proves the first claim.

Consider now the case where the intersection is $n - 2$. Suppose there exist three subspaces $A, B, C$ which intersect in a subspace of dimension $n - 2$. Let $D$ be another subspace. Then $D$ intersects $B$ in a subspace of dimension $n - 2$. Either this subspace is $A \cap B \cap C$, or not. We will suppose not, and show that this yields a contradiction. Then since $B$ and $D$ each are $(n - 1)$-dimensional, it must be that $B \cap D$ is formed by an $(n - 3)$-dimensional subspace of $A \cap B \cap C$, and a one-dimensional subspace that belongs to $B$, and not to $A$, or to $C$. Now $D$ also intersects $C$ in an $(n - 2)$ subspace, and it already intersects it in a $(n - 3)$-subspace, therefore the intersection of $C$ and $D$ further contains a one dimensional subspace that belongs to $C$, and not to $A$ and $B$. But by now, the $(n - 1)$-dimensional subspace $D$ is fully determined, it is made of a $(n - 3)$-dimensional subspace that belongs to $A \cap B \cap C$, and two other one-dimensional subspaces, one, say $c$, belonging to $C$ (and no to $A$ or $B$) and one, say $b$, belonging to $B$ (and not to $C$ or $A$). Finally, $D$ must intersect $A$ in a $(n - 2)$-dimensional subspace. So either $b$ or $c$ must belong to $A$. But this is not possible, because if $b$ belongs to $A$, then the intersection between $B$ and $A$ becomes of dimension $n - 1$, and similarly for if $c$ were to belong to $A$, then the intersection of $A$ and $C$ would be $n - 1$.

This shows that if there exist three subspaces that intersect in a subspace of dimension $n - 2$, then any other subspace will intersect them in the same subspace.
of dimension $n - 2$. Clearly this configuration cannot happen for the orbit code generated from a companion matrix. To see this, it is enough to consider $Ug, Ug^2$ and $Ug^3$ to see that $Ug^3$ contains a one-dimensional subspace which is common to $Ug^2$ but not to $Ug$.

We propose next a simple construction which will turn out to provide the optimal repair bandwidth, since it requires per failure the download of exactly the stored amount per node.

**Proposition 2.** Let $g$ be the companion matrix of $p(x)$ as in (4) and let $U_l$ contain a canonical basis of $U_l \in G_2(n - 1, n)$. Then the orbit code $UG$ has the following properties:

1. it is of length $l$ which is the order of $g$,
2. the object is retrieved by contacting any two nodes,

Furthermore, for $n \geq 4$, if $p(x)$ is chosen as in Corollary 1, then the repair of $t$ failures, $t \geq 2$, is done by downloading $n - 2$ elements of $\mathbb{F}_2$ from one node for each failure, and exchanging one element of $\mathbb{F}_2$ between two repair nodes.

**Proof.** The claim on the length comes from Lemma 4. Label the $l$ storage nodes from 0 to $l - 1$. By assumption, the $i$th node stores the element $Ug^i$ of the orbit $UG$ of $U$ under the right action of $G$, $i = 0, 1, ..., l - 1$.

To retrieve a stored object $o \in \mathbb{F}_2^{1 \times l}$ from any 2 nodes, or equivalently to protect against any $l - 2$ failures, we need that for a choice of 2 subspaces from the code $UG$,

$$\dim(U_{i_1} + U_{i_2}) = n,$$

and the claim follows from Lemma 5.

Recall Subsection 2.2 for counting the size of elements downloaded/repaired.

Consider any four distinct nodes say $s, v, i, j$, respectively storing $Ug^s, Ug^v, Ug^i, Ug^j$. Let $A = Ug^s \cap Ug^v \cap Ug^i \cap Ug^j$ be the intersection of node $i$ and $s$, it is of dimension $n - 2$ by the above lemma, and similarly, set $B = Ug^i \cap Ug^v$. We claim that $\dim(A + B) = n$. Since

$$\dim(A + B) = \dim(A) + \dim(B) - \dim(A \cap B) = 2(n - 2) - \dim(A \cap B),$$

we need to show that $\dim(A \cap B) = n - 4$. Write $Ug^i = \langle g_1, ..., g_{n-2}, g_{b_i} \rangle$, $Ug^j = \langle g_1, ..., g_{n-2}, g_{j_0} \rangle$. Then $A = Ug^s \cap Ug^v = \langle g_1, ..., g_{s0-1}, g_{s0+1}, ..., g_{n-2}, g_{s0} \rangle$, otherwise the nodes $i, j$ and $s$ would intersect in the same subspace of dimension $n - 2$, which is not possible by Lemma 6, and for the same reason $B = Ug^i \cap Ug^v = \langle g_1, ..., g_{b_i-1}, g_{b_i+1}, ..., g_{n-2}, g_{b_i} \rangle$. Without loss of generality, suppose that $s_0 < v_0$.

Then

$$A \cap B = \langle g_1, ..., g_{s0-1}, g_{s0+1}, ..., g_{v0-1}, g_{v0+1}, ..., g_{n-2} \rangle$$

which has dimension $n - 4$.

Suppose that $t$ nodes have failed, say node $w$, $w = 0, ..., t - 1$, which each downloads from node $w+t$ the subspace $Ug^w \cap Ug^{w+t}$ of dimension $n - 2$ (the labeling of the nodes is done for simplicity and without loss of generality). Each repaired node, after the download phase, is missing only one 1-dimensional subspace. By the above claim, any pair $A := Ug^w \cap Ug^{w+t}, B := Ug^w \cap Ug^{w+t}$ has the property that $\dim((Ug^w \cap Ug^{w+t}) = n)$, thus every repaired node can contact one other node to complete the repair process. Indeed, since $\dim(A + B) = n$, write the missing basis vector $a$ at node $w$ in a basis $\{v_1, ..., v_n\}$ of $A + B$ as $a = \sum_{i=1}^{n} a_i v_i$. 

---

**Advances in Mathematics of Communications** Volume 10, No. 1 (2016), 113–130
If $v_1, \ldots, v_l \in A$, ask the symbol $\sum_{i=l+1}^{n} a_i v_i$ from $B$ in node $w'$. Iterate this process for the missing basis vector at other nodes.

We remark that this code requires the download of $n-1$ symbols per failure, for a storage of $n-1$ symbols per node, thus is using the minimal possible bandwidth. It is a collaborative code for $t=2$, since two nodes are participating to the repair, and they are exchanging data among each other, and a partially collaborative code for $t \geq 3$, since then each repaired node only needs to communicate with one other node.

**Example 4.** Consider $G_2(3,4)$, and the subspace $U$, with basis

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Let $G = \langle g \rangle$ be the cyclic subgroup of $GL_4(\mathbb{F}_q)$ generated by

$$g = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

The order of $g$ is 5 and the elements of the orbit codes are explicitly given by

$$U, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

We store them in node 0 to 4. Assume that node 1 and node 4 failed. To repair node 1, download $\{(0010), (0001)\}$ from node 2, to repair node 4, download $\{(1000), (0100)\}$ from node 0, then during collaboration, the node repairing node 1 can compute (0011) and send it, and the node repairing node 4 sends (0100) in exchange. Note that the strategy is not unique. To repair node 1, download alternatively $\{(0001), (0110)\}$ from node 3, to repair node 4, get $\{(1100), (1111)\}$ from node 2, then the node repairing node 1 can get (0011) and the one repairing node 4 can get (0111) from the collaboration.

### 3.3. Code construction III: spreads.

In this section, we interpret a known family of storage codes [9] in the language of orbit codes, using the result from [14] that shows the connection between spreads and orbit codes.

Consider the finite fields $\mathbb{F}_q \subset \mathbb{F}_{q^n} \subset \mathbb{F}_{q^n}^*$, with $m \mid n$. It is well known that $\mathbb{F}_{q^n}$ is an $n$-dimensional vector space over $\mathbb{F}_q$, of which $\mathbb{F}_{q^m}$ and its cosets are $m$-dimensional subspaces. They thus form a spread, that is a collection of disjoint subspaces whose union is $\mathbb{F}_{q^n}$. Spreads to design storage codes have been investigated in [9], and can be obtained as a particular case of the proposed construction of orbit codes using companion matrices. For this, two ingredients are needed: (1) $g$ must be the companion matrix of a primitive polynomial $p(x)$, meaning that $p(x)$ is monic and irreducible of degree $n$, and of order $l = q^n - 1$, namely, roots of $p(x)$ are primitive elements of $\mathbb{F}_{q^n}$, they generate the cyclic group $\mathbb{F}_{q^n}^*$, and (2) $U$ is chosen so as to contain a basis for $\mathbb{F}_{q^m}$.

**Proposition 3.** Suppose that $m \mid n$, and consider the field extension $\mathbb{F}_{q^n} / \mathbb{F}_{q^m}$. Let $U$ contain a basis of $\mathbb{F}_{q^m}$, and $g$ be the companion matrix of a primitive polynomial $p(x)$ such that $p(\omega) = 0$ and $\langle \omega \rangle = \mathbb{F}_{q^n}^*$. Then $U\langle g \rangle$ is an orbit code of length $\frac{q^n-1}{q^m-1}$.
such that $Ug^i \cap Ug^j$ is trivial whenever $i \neq j$, and the union of $Ug^i$, $i = 1, \ldots, \frac{n}{m}$ is $F_q^n$.

Proof. Note that $\mu = \omega^{n/m}$ is a generator of $F_q^*$, and that $\{1, \mu, \ldots, \mu^{m-1}\}$ forms an $F_q$-basis of $F_q^m$. Form a matrix $U$ where each row contains $\mu^i = \omega^{in/m}$ written in an $F_q$-basis of $F_q^n$, $i = 0, \ldots, m - 1$. Then the action of $g$ corresponds to multiplication by $\omega$, and the claims follow from the fact that as a $F_q^*$ can be written as the union of $\omega^iF_q^m$, $i = 1, \ldots, \frac{q^n-1}{q^m-1}$.

More discussion about spread codes is also available in [14].

Example 5. Take $p(x) = x^4 + x + 1 \in F_2[x]$, it is a primitive polynomial, and let $g$ be its companion matrix

$$g = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

Let $\omega$ be a root of $p(x)$, that is $\omega^4 = \omega + 1$, also $\omega$ has order 15 and is a primitive element of $F_{16}$. Therefore $\omega^5$ is a generator of $F_4^*$, and an $F_2$-basis for $F_4$ is

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

Then the orbit code $U(g)$ forms a spread, it has length 5, and we obtain a known storage code [9, Example 1], namely

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

4. Instances of cyclic orbit codes using diagonal matrices

We describe in this section cyclic orbit codes based on diagonal matrices, and obtain Reed-Solomon codes as a particular case, and a family of storage codes which will turn out to give optimal parameters.

Let $\omega$ be the primitive element of $F_q$. Consider the cyclic group $G = \langle g \rangle$ of $GL_n(F_q)$ generated by the $n \times n$ diagonal matrix

$$g = \begin{bmatrix} 1 & & & \\ w & w^2 & & \\ & \ddots & \ddots & \\ & & w^{n-1} & \end{bmatrix}$$

where $n < |F_q|$. Then the order of $g$ is that of $w$, which is $|F_q| - 1$. Indeed, for $g^i$ to be the identity for some $i$, we at least need $w^i$ to be 1, but then all other powers $(w^j)^i$ will be 1 as well, $j = 2, \ldots, n - 1$.

4.1. Reed-Solomon codes. Consider the orbit code $UG$ where $U$ has for basis $U = [1, w, \ldots, w^{n-1}]$.

The length $N$ of the orbit code $UG$ is

$$N = |UG| = \frac{|G|}{|\text{Stab}_G(U)|} = |F_q| - 1.$$
In fact, the orbit code $UG$ is given by $UG = \{ Ug^i, \ i = 0, \ldots, |F_q| - 2 \}$ where

$$
Ug^i = \begin{bmatrix}
1 & w^i & w^{2i} & \cdots & w^{(n-1)i}
\end{bmatrix}
$$

$$
= \begin{bmatrix}
1, w^{i+1}, w^{2i+2}, \ldots, w^{(n-1)i+(n-1)}
\end{bmatrix}.
$$

This corresponds to a storage allocation where each node stores the inner product of the object $o$ with a column of a generator matrix of a Reed-Solomon code.

4.2. A code construction for $t$ failures, $t \geq 1$. A variation of the above construction can be obtained by tensor products. Consider the orbit code $UG$ where $G = \langle g \otimes I_m \rangle$ for $g$ the diagonal matrix given by

$$
g = \begin{bmatrix}
1 & w & w^2 & \cdots & w^{k-1}
\end{bmatrix},
$$

$k \geq 1$, and $I_m$ the $m$-dimensional identity matrix, and take for a basis of $U$

$$
U \otimes I_m = [1, w, \ldots, w^{k-1}] \otimes I_m.
$$

Note that the generator $g$ chosen for this construction is similar to (5), but for the dimension $k \geq 1$. Since $n$ is the size of the object, we choose $k$ and $m$ such that $km = n$. Since the order of $g \otimes I_m$ is still $|F_q| - 1$, this does not change the size $N$ of the orbit code $UG$.

**Proposition 4.** Let $g$ be the $k$-dimensional diagonal matrix above, and let $U \otimes I_m = [1, w, \ldots, w^{k-1}] \otimes I_m$ be a basis of the $m$-dimensional subspace $U \in G_q(m, n)$, $n = km$. Then the orbit code $UG$ has the following properties:

1. the object is retrieved by contacting any $k$ nodes,
2. the repair of $t$ failures, $t \geq 1$, is done by contacting any $k$ nodes.

**Proof.**

1. If any $k$ nodes $i_1, \ldots, i_k$ are contacted, then we obtain $k$ columns of a Vandermonde matrix, each tensored by $I_m$, thus a $k \times k$ invertible submatrix tensored by $I_m$, and the data object $o$ can be recovered.

2. Without loss of generality, we label the $t$ failed nodes from 1 to $t$. Then $t$ live nodes each contact $k$ nodes, say the $i$th node among those $t$ live nodes connects to nodes $i_1, \ldots, i_k$ and downloads only the one-dimensional subspace, corresponding to the $i$th row of $Ug^i \otimes I_m$, $j = 1, \ldots, k$. Again using the fact that we obtain $k$ columns of a Vandermonde matrix, the $i$th node computes a part of the object, given by the coefficients $o_i, o_{m+i}, \ldots, o_{(k-1)m+i}$, and after the download phase, it can compute the inner product of the vector $(o_i, o_{m+i}, \ldots, o_{(k-1)m+i})$ with the vector $Ug^i$, and thus obtains one of the $m$ coefficients it needs to store. The $m - 1$ missing coefficients are obtained through (partial) collaboration as follows: the $i$th node contacts nodes $i + 1, \ldots, i + m - 1$ (the indices are understood modulo the number of nodes). It already owns $(o_i, o_{m+i}, \ldots, o_{(k-1)m+i})(Ug^i)^T$, and gets $(o_{i+1}, o_{m+i+1}, \ldots, o_{(k-1)m+i+1})(Ug^i)^T$ from node $i + 1$, and similarly from
the other nodes, until \((o_i + m - 1, o_{m+i+m-1}, \ldots, o_{(k-1)m+i+m-1})(Ug^i)^T\) from node \(i + m - 1\).

5. Performance of storage codes

We briefly summarize known bounds on storage codes, to evaluate the proposed codes, and substantiate our earlier claim that they provide low repair bandwidth.

Recall that we encode an object \(o\) of size \(n\) over \(\mathbb{F}_q\) into \(N\) pieces and distribute them into \(N\) storage nodes, each of the nodes stores \(m\) coefficients in \(\mathbb{F}_q\).

5.1. Bounds. Denote by \(\alpha\) the storage capacity of every storage node, to keep the usual notation found in the literature, here, \(\alpha = m\). The typical known trade-off bounds for storage codes involve \(\alpha\), and the amount of data communicated during repair. They all assume that the storage codes have the property that any \(k\) out of \(N\) storage nodes contain enough information to recover the object \(o\).

1. If the system repairs one failure at a time, the communication cost in terms of bandwidth, called repair bandwidth, is \(\gamma = \frac{d\beta}{d}\) per node involved in the repair, where \(d\) is the number of live nodes contacted, and from each of them, \(\beta\) coefficients in \(\mathbb{F}_q\) are downloaded.

2. If the system waits for a threshold \(t\) of failures before starting the repair, then \(\gamma = \frac{d\beta + (t-1)\beta'}{d}\) per node, where \(t\) nodes each download \(\beta\) symbols of \(\mathbb{F}_q\) from \(d\) live nodes, and exchange \(\beta'\) symbols among each others. This scenario falls into the collaborative repair setting [2, 13].

3. The collaborative repair process can be refined so that the exchange phase of data does not involve all the repair nodes, which is called partially collaborative repair [5]. This is the most general setting as of now, and it subsumes the previous two cases. We introduce the parameter \(s\) to characterize the degree of collaboration, and \(t - s\) counts the number of repair nodes involved in the exchange process.

There are two extreme regimes when it comes to storage capacity and repair bandwidth: one minimizes the storage capacity, it is called the Minimum Storage Repair (MSR) point, while the other minimizes the repair bandwidth \(\gamma\) (this is Minimum Repair Bandwidth (MBR) point). At MSR point, as mentioned in Subsection 2.2, the storage capacity \(\alpha = m\) is \(n/k\): indeed, this follows from the assumption that any \(k\) nodes contain enough information to recover \(o\) of length \(n\). At MBR point, the minimum repair bandwidth \(\gamma\) per node is equal to \(\alpha\): this surely is the least amount of data to be transferred for a node to be repaired (see also Subsection 2.4).

What the trade-off computations between \(\alpha\) and \(\gamma\) tell [5] which is not obvious is that once \(\alpha = m = n/k\), \(\gamma\) cannot be arbitrarily low, and vice-versa, once \(\gamma = \alpha\), \(\alpha\) cannot be arbitrarily low. The lowest values that the different parameters achieve are in fact:

**MSR:** \(\alpha = \frac{n}{k}, \gamma = \frac{n}{k} \frac{d+t-s}{d-k-t+s+1}, \beta = \frac{n}{k} \frac{1}{d-k-t+s+1}\)

**MBR:** \(\alpha = \gamma = \frac{n}{k} \frac{2d+t-s}{2d-k-t+s+1}, \beta = \frac{n}{k} \frac{2}{2d-k-t+s+1}, \beta' = \frac{n}{k} \frac{1}{2d-k-t+s+1}\)

The meaning of the parameters involved is summarized for convenience in Table 5.1.
size of an object
any choice of \( k \) nodes allow the object retrieval
storage capacity per node
repair bandwidth per node
download repair bandwidth
exchange repair bandwidth
threshold at which a repair process is triggered
number of other repair nodes involved in collaboration

Table 1. Parameters involved in the bounds for storage codes.

When \( s = 1 \), we get the full collaboration case ((2) above), while \( s = t \) corresponds to the situation with no collaboration ((1) above).

5.2. Comparison. Let us consider the collaborative code construction of Subsection 3.2 for \( t = 2 \). An object \( o \) of length \( n \) is stored in \( N \) nodes, each of the nodes storing one codeword of the orbit code \( U_{i}(G) \), where \( U_{i} \in G_{2}(n-1, n) \) with canonical basis \( U_{i} \), thus \( \alpha = n - 1 \). The group \( G \) acting is \( G = \langle g \rangle \) with \( g \) the companion matrix of the polynomial \( p(x) \) whose order is \( l \), so that \( N = l \). When \( p(x) = \sum_{i=0}^{n} x^{i} \), we know from Proposition 2 that \( k = 2, \beta = n - 2, \beta' = 1 \). We notice that the total amount of data downloaded for repair is \( 2(n-2) + 2 = 2n - 2 \), which is \( n - 1 \) per node, suggesting that the code is at MBR. At MBR, we should have, using \( d = 1, k = 2, t - s = 1 \):

\[
\alpha = \gamma = \frac{3n}{4}, \ \beta = \frac{n}{2}, \ \beta' = \frac{n}{4}.
\]

When \( n = 4 \), the code in Subsection 3.2 has

\[
\alpha = \gamma = n - 1 = 3 = \frac{3n}{4}, \ \beta = n - 2 = 2 = \frac{n}{2}, \ \beta' = 1 = \frac{n}{4},
\]

which is indeed at MBR (\( \alpha \) is minimized). When \( n > 4 \), the code has parameters:

\[
\alpha = \gamma = n - 1 > \frac{3n}{4}, \ \beta = n - 2 > \frac{n}{2}, \ \beta' = 1 < \frac{n}{4},
\]

which, as mentioned above, satisfies \( \alpha = \gamma \), but \( \alpha \) is not minimized. The number of contacted nodes is optimal.

Let us consider the partially collaborative code construction of Subsection 3.2 for \( t \geq 3 \). As discussed above for the collaborative code construction, this partially collaborative code construction only changes the value of \( t \), but \( t - s \) still equals to 1. We know then that the repair bandwidth is still \( n - 1 \), which suggests that the code is at MBR. Since the optimal values of \( \alpha, \gamma \) at MBR point are related to the value of \( t - s \) which is not changed, the minimized \( \alpha \) is obtained by \( n = 4 \) (which indeed is a collaborative phase since three failures cannot happen when \( n = 4 \)). For \( n \geq 4 \), still \( \alpha = \gamma \), but \( \alpha \) is not minimized, while the number of contacted nodes stays optimal. The gap between the optimal storage and the value of \( \alpha \) for this construction is shown on Figure 1, for different values of \( n \) between 4 and 10. We remark that optimal code constructions at MBR for partial collaboration are not known as of now. This suggests that a deeper study of automorphism groups as a future work to get new/better codes for partial collaboration.

Consider finally the code given in Subsection 4.2. Set \( m \) to be \( t - s + 1 \), which sets the level of collaboration, namely there is collaboration among \( t - s \) repair nodes. Then the size of the object is \( n = k(t - s + 1) \), each node has a storage capacity of
The storage code proposed in Subsection 3 is optimal with respect to the minimum repair bandwidth, but not with respect to the storage capacity $\alpha$. This shows the optimal value of $\alpha$ with respect to the storage needed for the construction, as a function of $n$. When $n = 4$, our construction is optimal. We set $d = 1$, $k = 2$ and $t - s = 1$.

$\alpha = t - s + 1, t - s \geq 1$. This construction can be seen as a new interpretation of the construction of [6] which fits the minimum storage regime (MSR), since $\alpha = n/k$. The total repair bandwidth $\gamma$ is

$$\gamma = k + (t - s), \quad 1 \leq s \leq t,$$

since every repair node downloads $k$ amount of data, and exchanges $t - s$ of it.

**REFERENCES**

[1] H. D. L. Hollmann, Storage codes – coding rate and repair locality, in *Int. Conf. Comp. Netw. Commun.*, 2013, 830–834.

[2] A.-M. Kermarrec, N. Le Scouarnec and G. Straub, Repairing multiple failures with coordinated and adaptive regenerating codes, in *Int. Symp. Network Coding*, IEEE, 2011, 1–6.

[3] S. Kim, Linear algebra from module theory perspective, available online at [http://www.math.ucla.edu/~i707107/LinearAlgebra.pdf](http://www.math.ucla.edu/~i707107/LinearAlgebra.pdf)

[4] R. Lidl and H. Niederreiter, *Introduction to Finite Fields and their Applications*, Cambridge Univ. Press, Cambridge, 1986.

[5] S. Liu and F. Oggier, On storage codes allowing partially collaborative repairs, in *IEEE Int. Symp. Inf. Theory*, 2014, 2440–2444.

[6] S. Liu and F. Oggier, Two storage code constructions allowing partially collaborative repairs, in *Int. Symp. Inf. Theory Appl.*, IEEE, 2014, 378–382.

[7] S. Liu and F. Oggier, On the Design of Storage Orbit Codes, in *Coding Theory and Applications*, Springer, 2015, 263–271.

[8] F. Oggier, Some constructions of storage codes from Grassmann graphs, in *Proc. Int. Zurich Sem. Commun.*, Zürich, 2014.

[9] F. Oggier and A. Datta, Self-repairing codes for distributed storage - a projective geometric construction, in *IEEE Inf. Theory Workshop*, 2011, 30–34.

[10] F. Oggier and A. Datta, *Coding Techniques for Repairability in Networked Distributed Storage Systems*, Now Publishers, 2013.

[11] K. V. Rashmi, N. B. Shah, P. V. Kumar and K. Ramchandran, Explicit construction of optimal exact regenerating codes for distributed storage, in *47th Ann. Allerton Conf. Commun. Control Comput.*, IEEE, 2009, 1243–1249.
[12] N. Raviv and T. Etzion, Distributed storage systems based on equidistant subspace codes, preprint, arXiv:1406.6170

[13] K. W. Shum and Y. Hu, Cooperative regenerating codes, IEEE Trans. Inf. Theory, 59 (2013), 7229–7258.

[14] A.-L. Trautmann, F. Manganiello, M. Braun and J. Rosenthal, Cyclic orbit codes, IEEE Trans. Inf. Theory, 59 (2013), 7386–7404.

[15] Z. Wang, I. Tamo and J. Bruck, Long MDS codes for optimal repair bandwidth, in IEEE Int. Symp. Inf. Theory, 2012, 1182–1186.

Received December 2014; revised July 2015.

E-mail address: SLIU012@e.ntu.edu.sg
E-mail address: frederique@ntu.edu.sg