Coding over random networks

Ivo Kubjas

Introduction

This report is based on the article [1] by Koetter and Kschischang.

The authors introduced a novel approach for constructing codes by selecting a subset of a powerset of a vector space. With the code construction, they also defined an operator channel for code transmission. The codewords are vector subspaces. The data is transmitted by using one of the bases of such a subspace. The properties of codes of this type were studied.

In this report, an overview of the operator channel, the codes and different bounds on the code’s error correction capabilities are given.

1 Preliminaries

The article makes extensive use of vector spaces and their subspaces. For completeness, we give a short introduction to basic aspects of the theory.

Definition 1. Let $W$ be a vector space. We call a collection of all subspaces of $W$ a powerset of $W$. The powerset of $W$ is denoted as $\mathcal{P}(W)$.

As empty set $\emptyset = \{\}$ and $W$ itself are subspaces of $W$, then $\emptyset \in \mathcal{P}(W)$ and $W \in \mathcal{P}(W)$.

Definition 2. Let $W$ be a vector space over the field $\mathbb{F}$ and let $W_\alpha, \alpha \in I$, be subspaces of $W$. We call the space

$$\sum_{\alpha \in I} W_\alpha = \{x \in W : x = \sum_{\alpha \in I} x_\alpha, x_\alpha \in W_\alpha\}$$

a sum of spaces $W_\alpha$.

If $I = \{1, \ldots, n\}$, then we can also write the sum of spaces $W_i$ as

$$W_1 + W_2 + \cdots + W_n.$$ 

If the subspaces $W_\alpha, \alpha \in I$, are pairwise disjoint, then the sum of vector spaces is called the direct sum of these vector spaces and in finite case is written as

$$W_1 \oplus W_2 \oplus \cdots \oplus W_n.$$
**Definition 3.** The span of a finite subset $S$ of a vector space $W$ over the field $F$ is a set, which consists of all linear combinations of vectors in $S$.

The span of a set $S$ is denoted as $\text{span}(S) = \langle S \rangle$. The dimension of a vector space is the size of the smallest set of linearly independent elements which spans the vector space.

In the context of this summary, we are interested in finite spans. It can be shown that the span is also a vector space.

## 2 Random networks

We consider the case where the network consists of arbitrary number of nodes and edges between the nodes.

![Diagram](image)

**Figure 1:** Example of transmitting of a vector space $V = \langle \{e_1, e_2\} \rangle$ in a network. Receiver spans vector space $U = \langle \{\gamma_{11}, \gamma_{12}\} \rangle$ from received vectors $\gamma_{11}$ and $\gamma_{12}$.

If $I$ is the set of indexes of the basis vectors, then consider packets $p_i \in F_q^N (i \in I)$ from the set of the basis vectors of an $N$-dimensional vector subspace $V$ over the field $F_q$.

The transmission is split into generations, during which the sender injects packets into the network. Intermediate nodes receive the packets and send a random linear combination of received packets.
In the error-free case this means that the receiver receives the packets
\[ y_j = \sum_{i=1}^{M} h_{j,i} p_i \]
where \( p_i \) are injected packets and \( h_{j,i} \) are the random coefficients in \( \mathbb{F}_q \) for the \( j \)-th received packet.

If error packets (e.g., from communication errors) are injected into the network, then the receiver receives
\[
y_j = \sum_{i=1}^{M} h_{j,i} p_i + \sum_{t=1}^{T} g_{j,t} e_t, \tag{1}
\]
where \( e_t, \ t = 1, \ldots, T \) are error packets. Here, \( g_{j,t} \in \mathbb{F}_q \) are coefficients which describe the errors and are unknown to the receiver. If an error packet is injected by a node, then consecutive nodes keep the error and possibly transmit erroneous packets further. In the most extreme case, a single error packet may lead to errors in every received packet.

We can write (1) in the matrix form as
\[
y = Hp + Ge \tag{2}
\]
Here, \( H \) and \( G \) are matrices of size \( L \times M \) and \( L \times T \), \( p \) is the \( M \times N \) matrix of transmitted vectors and \( e \) is the \( T \times N \) error matrix. The received vectors form a \( L \times N \) matrix \( y \).

If we omit the error then the received packets will be the row vectors from \( y = Hp \). As the intermediate nodes send random linear combinations of received packets, then \( H \) is random for the receiver. If we consider full rank matrices \( H \), then there are no restrictions on \( p \). The receiver can however generate the vector space span by the row vectors of \( Hp \).

Instead of using a collection of packets as the messages (as in the classical setting), we consider a vector space as a message to be transmitted. More generally, let \( W \) be a \( N \)-dimensional vector space over \( \mathbb{F}_q \). The message will be a subspace \( V \) of \( W \) and it will be transmitted using its basis elements.

An example transmission is illustrated in Figure 1.

The dimension of the element \( V \in \mathcal{P}(W) \) is denoted by \( \dim(V) \). The sum of two subspaces is denoted by \( V + U = \{ x + y : x \in V, y \in U \} \). If two subspaces have trivial intersection, then their sum \( V + U \) is their direct sum \( V \oplus U \) and in this case \( \dim(V + U) = \dim(V) + \dim(U) \). Also, for any subspaces \( V \) and \( U \) we have \( V = (V \cap U) \oplus V' \) for some subspace \( V' \) isomorphic to the quotient space \( V/(V \cap U) \). As \( (V \cap U) \in U \), then \( U + V = U + ((V \cap U) \oplus V') = U \oplus V' \).

2.1 Operator channel

We give a more formal definition for an operator channel. Firstly, we define an operator \( \mathcal{H}_k : \mathcal{P}(W) \to \mathcal{P}(W) \) which selects a random subspace from
given vector space. If \( \dim(V) \geq k \), then it returns a random \( k \)-dimensional subspace of \( V \) and otherwise it returns \( V \) itself. We can intuitively look at \( \mathcal{H}_k \) as being an erasure operator, as it removes some information about initial space \( V \).

There are no further assumptions about \( \mathcal{H}_k \) and \( \mathcal{H}_k(V) \) can be arbitrary. In addition to erasure operator we also define error operator as a vector space \( E \), which introduces errors into the network. Erasure and error operator differ as the former loses information about the initial vector space \( V \) and the error operator adds additional (false) information to the vector space.

Using these operators we can give a related definition to \( \mathcal{H}_k \).

**Definition 4.** An operator channel \( C \) associated with the ambient space \( W \) is a channel with input and output alphabet \( \mathcal{P}(W) \). The channel input \( V \) and channel output \( U \) is related as

\[
U = \mathcal{H}_k(V) \oplus E,
\]

where \( k = \dim(U \cap V) \) and \( E \) is an error operator. We say that the operator channel commits \( \rho = \dim(V) - k \) erasures and \( t = \dim(E) \) errors.

### 2.2 A metric on \( \mathcal{P}(W) \)

The metric on \( \mathcal{P}(W) \) is defined as a function \( d : \mathcal{P}(W) \times \mathcal{P}(W) \to \mathbb{Z}_+ \) where

\[
d(A, B) := \dim(A + B) - \dim(A \cap B).
\]

In [1], the following underlying result was proven:

**Lemma 1.** The function

\[
d(A, B) := \dim(A + B) - \dim(A \cap B)
\]

is a metric for the space \( \mathcal{P}(W) \).

The inner product for the elements in \( N \)-dimensional vector space \( W \) is defined as

\[
(u, v) := \sum_{i=1}^{N} u_i v_i,
\]

where \( u = (u_1, \ldots, u_N) \) and \( v = (v_1, \ldots, v_N) \) are represented by coordinates of the basis vectors of \( W \). Note, that the coordinates \( u_i, v_i \) are in \( \mathbb{F}_q \).

Having an inner product, we can also define orthogonal subspace \( U^\perp \) for \( U \in \mathcal{P}(W) \) as

\[
U^\perp := \{ v \in W : (u, v) = 0 \ \forall u \in U \}.
\]
We can see that if $U$ is an $k$-dimensional subspace then its orthogonal space $U^\perp$ is of dimension $N - k$. Furthermore:

\[
\begin{align*}
    d\left(U^\perp, V^\perp\right) &= \dim\left(U^\perp + V^\perp\right) - \dim\left(U^\perp \cap V^\perp\right) \\
    &= \dim\left((U \cap V)^\perp\right) - \dim\left((U + V)^\perp\right) \\
    &= (N - \dim(U \cap V)) - (N - \dim(U + V)) \\
    &= \dim(U + V) - \dim(U \cap V) \\
    &= d(U, V)
\end{align*}
\]

3 Codes in $\mathcal{P}(W)$

We have established necessary definitions for defining a code in $\mathcal{P}(W)$. The code $\mathcal{C}$ is a nonempty subset of $\mathcal{P}(W)$.

The minimum distance for the code $\mathcal{C}$ is defined as the minimum distance between elements in $\mathcal{C}$:

\[
D(\mathcal{C}) := \min_{X,Y \in \mathcal{C}, X \neq Y} d(X,Y),
\]

and the maximum dimension of the code $\mathcal{C}$ is defined as

\[
l(\mathcal{C}) := \max_{X \in \mathcal{C}} \dim(X).
\]

We say that the code $\mathcal{C}$ is $[N, l(\mathcal{C}), \log_q|\mathcal{C}|, D(\mathcal{C})]$-code if $W$ is $N$-dimensional space over $\mathbb{F}_q$, the minimum distance of the code is $D(\mathcal{C})$ and the maximum dimension is $l(\mathcal{C})$. A code has also normalized parameters.

**Definition 5.** Let $\mathcal{C}$ be a code of type $[N, l(\mathcal{C}), \log_q|\mathcal{C}|, D(\mathcal{C})]$. The normalized weight $\lambda$, the rate $R$ and the normalized minimum distance $\delta$ of $\mathcal{C}$ are defined as

\[
\lambda = \frac{l(\mathcal{C})}{N}, \quad R = \frac{\log_q|\mathcal{C}|}{NI(\mathcal{C})}, \quad \delta = \frac{D(\mathcal{C})}{2l(\mathcal{C})}
\]

The normalized weight $\lambda$ describes the ratio of codewords to the total number of subspaces. To transmit a vector $V \in \mathcal{C}$ up to $NI(\mathcal{C})$ symbols needs to be sent and the rate shows how efficient the code is. The normalized minimum distance $\delta$ describes error-correcting capabilities of the code. We later see that the normalized minimum distance and rate are related, i.e. higher rate $R$ leads to smaller minimum distance $\delta$ and this affects the error detection and correction capabilities of the code.

Naturally, one is interested in the error and erasure correcting capabilities of the code. A bound can be given for a minimum distance decoder.
We call a decoder, which takes the output $U$ of the operator channel and returns a codeword $V \in C$ such that for all $V' \in C$ we have $d(U, V) \leq d(U, V')$, a minimum distance decoder.

We define $(x)_+ := \max\{0, x\}$, positive part of $x$.

**Theorem 1.** Assume a code $C$ with parameters $[N, l, \log_q |C|, D]$ is used for transmission over an operator channel. Let $V \in C$ be transmitted and let

$$U = \mathcal{H}_k(V) \oplus E$$

be received. If $\dim(E) = t$ and $\rho = (l(C) - k)_+$ denote the maximum number of erasures induced by the channel and if

$$2(t + \rho) < D(C)$$

then a minimum distance decoder for $C$ will output the transmitted space $V$.

**Proof.** Let $V' = \mathcal{H}_k(V)$. From the triangle inequality we have $d(V, U) \leq d(V, V') + d(V', U) \leq \rho + t$. If $T \neq V$ is another codeword in $C$, then $D(C) \leq d(V, T) \leq d(V, U) + d(U, T)$, from which $d(U, T) \geq D(C) - \rho + t > \rho + t$. Then $d(U, T) > d(V, U)$, meaning that minimum distance decoder must produce $V$ as the codeword. \hfill \Box

If looking at two distinct cases, where the operator channel produces no erasures (by taking the operator $\mathcal{H}_{\dim(W)}$ as an identity) or when the channel produces no errors (by taking $E = \{0\}$), we achieve the following conclusion.

**Corollary 1.** Assume a code $C$ is used for transmission over an operator channel. Let $V \in C$ be transmitted and let

$$U = \mathcal{H}_{\dim(W)}(V) \oplus E = V \oplus E$$

be received. If $\dim(E) = t$ and $2t < D(C)$, then a minimum distance decoder for $C$ will produce the transmitted space $V$.

Symmetrically, let

$$U = \mathcal{H}_k(V) \oplus \{0\} = \mathcal{H}_k(V)$$

be received. If $2\rho < D(C)$ where $\rho = (l(C) - k)_+$, then a minimum distance decoder for $C$ will output the transmitted space $V$.

The corollary allows to achieve familiar bound on the dimension of correctable errors as

$$t \leq \left\lfloor \frac{D(C) - 1}{2} \right\rfloor.$$
3.1 Constant dimension codes

Setting restrictions on codes can have a side-effect of allowing the receiver to use knowledge about restrictions for building more efficient decoder.

If we have a \( l \)-dimensional \([N, l(C), M, D(C)]\)-code \( C \), then \( l(C) = l \). \( \text{w.l.o.g.} \) we can assume that \( l \leq N - l \). If \( l > N - l \) then we can define a complementary code \( C^\perp = \{ V^\perp : V \in C \} \), where the dimension of each codeword is \( l' = N - l \). We saw previously, that all distance properties were the same for complements of codewords (most importantly, \( D(C) = D(C^\perp) \)). As also \( l' \leq N - l' \), then we have a required code with the same properties as formerly.

Definition 6. The set of all subspaces of \( W \) of dimension \( l \) is denoted as \( \mathcal{P}(W, l) \).

Generally, \( \mathcal{P}(W, l) \) is also known as a Grassmannian. The Grassmann graph \( G_{W,l} \) has vertices from \( \mathcal{P}(W, l) \) with an edge between vertices \( V \) and \( U \) only if \( d(U,V) = 2 \).

If the dimension of the code is fixed, then the receiver collects a number of packets until it is able to recover \( l \)-dimensional space. If it collects more packets, then the possible outcome may be inconsistent (e.g. the dimension is larger than \( l \), clearly not belonging to \( \mathcal{P}(W, l) \)). In this case, the number of erasures and errors is equal and the corresponding operator channel outputs

\[
U = \mathcal{H}_{l-t(E)}(V) \oplus E
\]
on input \( V \), where \( t(E) \) is the number of erasures and errors.

If the dimension of the erasures and the dimension of the errors is the same (as in the code where dimension of vector spaces is fixed), then from Theorem 1 the minimum distance decoder can correct any combination of dimensional errors of size up to \( t(E) \leq \left\lfloor \frac{D(C) - 1}{4} \right\rfloor \).

4 Bounds on codes

4.1 \( q \)-ary Gaussian coefficients

The \( q \)-ary Gaussian coefficients provide an intuitive analogue to binomial coefficients.

Definition 7. \( q \)-ary Gaussian coefficient for non-negative integers \( l \) and \( n \) with \( l \leq n \) is

\[
\binom{n}{l}_q = \frac{(q^n - 1)(q^{n-1} - 1) \cdots (q^{n-l+1} - 1)}{(q^l - 1)(q^{l-1} - 1) \cdots (q - 1)} = \prod_{i=1}^{l-1} \frac{q^{n-i} - 1}{q^i - 1}
\]

If \( l = 0 \), then \( \binom{n}{0}_q \) is defined as \( \binom{n}{0}_q = 1 \).
Similarly to binomial coefficients, we can have

\[
\binom{n}{l}_q = \frac{\prod_{i=0}^{l-1} q^{n-i} - 1}{\prod_{i=0}^{l-1} q^{l-i} - 1} = \frac{\prod_{i=0}^{n-1} q^{n-i} - 1}{\prod_{i=0}^{l-1} q^{l-i} - 1}
\]

\[
= \frac{n-1}{\prod_{i=0}^{l-1} q^{n-l-i} - 1} \frac{\prod_{i=0}^{n-1} q^{n-i} - 1}{\prod_{i=0}^{n-l} q^{n-i} - 1} = \binom{n}{n-l}_q
\]

The Gaussian coefficients gives the number of \( l \)-dimensional subspaces of \( n \)-dimensional space over \( \mathbb{F}_q \). Thus, bounds on the Gaussian coefficients can help in estimating the size of the code.

**Lemma 2.** If \( 0 < l < n \), then the Gaussian coefficient is bounded by

\[1 < q^{-l(n-l)} \binom{n}{l}_q < 4.\]

**Proof.** Considering matrices of the form \([I|A]\) over \( \mathbb{F}_q \), where \( I \) is an \( l \times l \) identity matrix and \( A \) is an arbitrary \( l \times (n-l) \) matrix, then there are in total of \( q^{l(n-l)} \) such matrices.

Row space of each of this matrix spans a \( l \)-dimensional vector subspace in \( \mathbb{F}_q^n \). As \( l > 0 \), then not all \( l \)-dimensional vector subspaces of \( \mathbb{F}_q^n \) are generated. Thus

\[q^{l(n-l)} < \binom{n}{l}_q,
\]

and lower bound is achieved.

For the upper bound,

\[
\binom{n}{l}_q = q^{l(n-l)} \frac{(1-q^{-n})(1-q^{-n+1}) \cdots (1-q^{-n+l-1})}{(1-q^{-l})(1-q^{-l+1}) \cdots (1-q^{-1})}
\]

\[
< q^{l(n-l)} \frac{1}{(1-q^{-l})(1-q^{-l+1}) \cdots (1-q^{-1})}
\]

\[
< q^{l(n-l)} \prod_{j=1}^{\infty} \frac{1}{1-q^{-j}}
\]

It is known from analytical combinatorics, that \( f(x) = \prod_{j=1}^{\infty} \frac{1}{1-x^j} \) is the generating function for integer partitions. As \( q \geq 2 \), then

\[
\prod_{j=1}^{\infty} \frac{1}{1-q^{-j}} \leq \prod_{j=1}^{\infty} \frac{1}{1-2^{-j}} = \frac{1}{Q_0} < 4,
\]

where \( Q_0 \approx 0.289 \) is a constant, which gives the probability that large randomly chosen matrix over \( \mathbb{F}_2 \) is nonsingular. \( \square \)
4.2 Spheres in subspaces

As the code has an associated metric measuring distance between the codewords, it is possible to define a sphere around each codeword.

**Definition 8.** Let $W$ be an $N$-dimensional vector space and let $\mathcal{P}(W, l)$ be the set of $l$-dimensional subspaces of $W$. The sphere $S(V, l, t)$ of radius $t$ centered at space $V$ in $\mathcal{P}(W, l)$ is defined as the set of all subspaces $U$ that satisfy $d(U, V) \leq 2t$. Symbolically,

$$S(V, l, t) = \{ U \in \mathcal{P}(W, l) : d(U, V) \leq 2t \}.$$ 

We note that radius $2t$ is used because codewords in $\mathcal{P}(W, l)$ form a Grassmann graph, thus the distance between two codewords is twice the distance of the codewords in the Grassmann graph.

**Theorem 2.** The number of spaces in $S(V, l, t)$ does not depend on $V$ and equals to

$$|S(V, l, t)| = \sum_{i=0}^{t} q^{2 \left[ \begin{array}{c} l \\ i \\ q \end{array} \right] \left[ \begin{array}{c} N - l \\ i \\ q \end{array} \right]}$$

for $t \leq l$.

**Proof.** We look at the spaces $U$ which intersect $V$ in $l - i$ dimensional subspace. The intersection subspace can be chosen in $\left[ \begin{array}{c} l \\ l - i \\ q \end{array} \right] = \left[ \begin{array}{c} l \\ i \\ q \end{array} \right]$ ways.

There are

$$\frac{(q^N - q^l)(q^N - q^{l+1}) \cdots (q^N - q^{l+i-1})}{(q^l - q^{l-i})(q^l - q^{l-i+1}) \cdots (q^l - q^{l-1})} = \frac{q^{li}q^{i-1} \cdots q^{l}}{q^{l-i}q^{N-l-i-1} \cdots q^{N-l+i+1} \left[ \begin{array}{c} N - l \\ i \\ q \end{array} \right]}$$

$$= q^{(N-i)-i} \sum_{j=0}^{l-i-j-N+l+j} \left[ \begin{array}{c} N - l \\ i \\ q \end{array} \right]$$

$$= q^{\sum_{j=0}^{l-i} \left[ \begin{array}{c} N - l \\ i \\ q \end{array} \right]} = q^{2 \left[ \begin{array}{c} N - l \\ i \\ q \end{array} \right]}$$

ways to complete the remaining part of the intersection.

Therefore, the number of codewords at distance $2t$ from $V$ is $q^{2 \left[ \begin{array}{c} N - l \\ i \\ q \end{array} \right]} \left[ \begin{array}{c} l \\ i \\ q \end{array} \right]$. If we sum this expression over all $i \leq t$, then we get the desired result.

4.3 Sphere-packing and sphere-covering bounds

The sphere-packing upper bound follows directly from the results of previous sections.
Theorem 3. Let \( \mathcal{C} \) be a collection of spaces in \( \mathcal{P}(W, l) \) such that \( D(\mathcal{C}) \geq 2t \) and let \( s = \lfloor \frac{t-1}{2} \rfloor \). The size of \( \mathcal{C} \) must satisfy

\[
|\mathcal{C}| \leq \frac{|\mathcal{P}(W, l)|}{|S(V, l, s)|} = \frac{n}{l} \frac{N}{t} q^{N - l} \frac{q}{s} q < 4q^{(l-s)(N-s-l)}.
\]

Proof. We use Theorem 2 and Lemma 2. 

Similarly, we can easily obtain the sphere-covering lower bound.

Theorem 4. There exists a code \( \mathcal{C} \) with distance \( D(\mathcal{C}) \geq 2t \) such that \( |\mathcal{C}| \) is lower bounded by

\[
|\mathcal{C}| > \frac{N}{l} > \frac{q^{(l-t+1)(N-l-t+1)}}{(t-1)q^{t-1} \frac{l}{t} q^{N-l} \frac{1}{t-1} q}. 
\]

The bounds can be compared if the parameters are normalized.

Corollary 2. Let \( \mathcal{C} \) be a collection of spaces in \( \mathcal{P}(W, l) \) such that \( \delta = \frac{D(\mathcal{C})}{2l} \). The rate of \( \mathcal{C} \) is upper bounded by

\[
R \leq 1 - \frac{\delta}{2} (1 - \lambda (1 + \frac{\delta}{2})) + o(1).
\]

Proof.

\[
R = \frac{\log_q(|\mathcal{C}|)}{Nl} \leq \frac{(l-s)(N-s-l)}{Nl} + o(1)
\]

\[
\leq \left(1 - \frac{D(\mathcal{C})}{4l}\right) \left(1 - \frac{D(\mathcal{C})}{4N} - \frac{l}{N}\right) + o(1)
\]

\[
\leq \left(1 - \frac{\delta}{2}\right) \left(1 - \frac{\delta \lambda}{2} - \lambda\right) + o(1)
\]

\[
\leq \left(1 - \frac{\delta}{2}\right) \left(1 - \lambda \left(\frac{\delta}{2} + 1\right)\right) + o(1)
\]

Corollary 3. There exists a code \( \mathcal{C} \) such that the rate of the code of the code is lower bounded by

\[
R \geq (1 - \delta)(1 - \lambda(\delta + 1)) + o(1).
\]

Proof to this corollary is omitted due to its technicality.
4.4 Singleton bound

To derive the Singleton bound, a concept of puncturing is introduced. Puncturing allows to replace the initial code $C$ with a smaller code while maintaining the cardinality of the code.

Let $C$ be a collection of sets from $\mathcal{P}(W, l)$, where $W$ is a $N$-dimensional space. Let $W'$ be a random $(N-1)$-dimensional subspace of $W$. A code puncturing replaces each subspace $V \in C$ with a subspace $V' = \mathcal{H}_{l-1}(V \cap W')$. We denote the punctured code by $C|_{W'}$. Intuitively, puncturing of the code decreases the dimension of each subspace $C$ which is not contained in $W'$ by one.

Clearly, the maximum distance of the spaces in $C|_{W'}$ is $l-1$ and the dimension of $W$ is $N-1$. The following theorem shows the remaining parameters of the resulting code.

**Theorem 5.** If $C \subseteq \mathcal{P}(W, l)$ is a code of type $[N, l, \log_q(|C|), D]$ with $D > 2$, $W'$ is $N-1$-dimensional subspace of $W$, then $C|_{W'}$ is a code of type $[N-1, l-1, \log_q(|C|), D']$ with $D' \geq D - 2$.

**Proof.** Let $U' = \mathcal{H}_{l-1}(U \cap W')$ and $V' = \mathcal{H}_{l-1}(V \cap W')$ be codewords in $C|_{W'}$. From the definition of the distance we have $d(U, V) = 2l - 2 \dim(U \cap V) \geq D$ and so $2 \dim(U' \cap V') \leq 2 \dim(U \cap V) \leq 2l - D$. Applying the definition of the distance leads directly to $d(U', V') \geq D - 2$. Since $D > 2$, for any $U'$ and $V'$ we have $d(U', V') > 0$ and therefore no subspaces collide. Thus the cardinality of $C|_{W'}$ is the same as of the initial code $C$. \hfill $\square$

Direct consequence of the theorem is the main result of this subsection, the Singleton bound.

**Theorem 6.** A $q$-ary code $C \subseteq \mathcal{P}(W, l)$ of type $[N, l, \log_q(|C|), D]$ must satisfy

$$|C| \leq \left\lfloor \frac{N - (D - 2)/2}{\max \left( l, N - l \right)} \right\rfloor_q.$$

**Proof.** Puncturing the code $C$ for $(D - 2)/2$ times, the code $C'$ of type $[N - (D - 2)/2, l - (D - 2)/2, \log_q(|C|), D']$ with $D' \geq 2$ is obtained. This code has at most

$$\left\lfloor \frac{N - (D - 2)/2}{l - (D - 2)/2} \right\rfloor_q = \left\lfloor \frac{N - (D - 2)/2}{N - l} \right\rfloor_q$$

codewords.

Similarly, after puncturing code $C^\perp$, it has at most

$$\left\lfloor \frac{N - (D - 2)/2}{l} \right\rfloor_q$$

codewords.

It was shown that codes $C$ and $C^\perp$ are equivalent, thus $C$ must have the largest number of codewords of the two and the bound from the theorem follows. \hfill $\square$

The bound using normalized parameters is given by the following lemma.
Lemma 3. Let $C$ be a collection of spaces in $\mathcal{P}(W,l)$ with $l \leq \dim(W)/2$. The rate of $C$ is bounded from above by

$$R \leq (1 - \delta)(1 - \lambda) + \frac{1}{\lambda N}(1 - \lambda + o(1)).$$

Proof. If $l \leq \dim(W)/2$, then $N - l \geq l$. From the previous theorem

$$R = \frac{\log_q |C|}{Nl} \leq \sum_{i=0}^{N-l-1} \left( \frac{(N - (D - 2)/2 - i) - (N - L - i)}{Nl} \right) + o(1)$$

$$\leq \frac{(N - l)(l - (D - 2)/2)}{Nl} + o(1) \leq \left( 1 - \frac{l}{N} \right) \left( 1 - \frac{D}{2l} \right) + \frac{1}{l} - \frac{1}{N} + o(1)$$

$$\leq (1 - \lambda)(1 - \delta) + \frac{1}{\lambda N}(1 - \lambda + o(1))$$

Figure 2: Comparison of different bounds

Comparison of the bounds can be seen on Figure 2.
5 Code construction

5.1 Linearized polynomials

The code construction makes heavy use of linearized polynomials. A linearized polynomial is a polynomial over \( \mathbb{F} = \mathbb{F}^n_q \), which has the following form:

\[
L(x) = \sum_{i=0}^{d} a_i x^q^i.
\]

If \( q \) is clear from the context, then we can also write \([i] = q^i\).

\( \mathbb{F} \)-linear combination of linearized polynomials is a polynomial which has the form

\[
\alpha_1 L_1(x) + \alpha_2 L_2(x),
\]

where \( \alpha_1, \alpha_2 \in \mathbb{F} \) and \( L_1(x) \) and \( L_2(x) \) are linearized polynomials.

Composition of linearized polynomials is

\[
L_1(x) \otimes L_2(x) = L_1(L_2(x)).
\]

It can be easily shown that linear combination of linearized polynomials is linearized polynomial and that the composition \( \oplus \) of linearized polynomials is also a linearized polynomial, thus linearized polynomials form a non-commutative ring under addition and composition.

Linearized polynomial is said to be a zero polynomial if all the coefficients \( a_i \) are zero. Zero polynomial is denoted as \( L(x) \equiv 0 \). Two linearized polynomials \( L_1(x) \) and \( L_2(x) \) are equivalent if \( L_1(x) - L_2(x) \equiv 0 \).

Linearized polynomials act as linear transformations. Let \( L(x) \) be a linearized polynomial over \( \mathbb{F} = \mathbb{F}_q^n \) and let \( K \) be an extension field for \( \mathbb{F} \). Then \( L(x) \) is linear in respect to \( \mathbb{F}_q \):

\[
L(\lambda_1 \alpha_1 + \lambda_2 \alpha_2) = \lambda_1 L(\alpha_1) + \lambda_2 L(\alpha_2),
\]

where \( \lambda_1, \lambda_2 \in \mathbb{F}_q \) and \( \alpha_1, \alpha_2 \in K \).

The equivalence of two linearized polynomials can also be checked by comparing the roots of the linearized polynomial.

**Theorem 7.** Let \( \mathbb{F} = \mathbb{F}_q^n \) be a extension field of \( \mathbb{F}_q \) and \( K \) be an extension field of \( \mathbb{F} \). Let \( d \) be a positive integer and let \( L_1(x) \) and \( L_2(x) \) be two linearized polynomials over \( \mathbb{F} \) of degree less than \( [d] \). If \( \alpha_1, \alpha_2, \ldots, \alpha_d \) are linearly independent elements of \( K \) such that \( L_1(\alpha_i) = L_2(\alpha_i) \) for \( i \in \{1, \ldots, d\} \) then \( L_1(x) \equiv L_2(x) \).

**Proof.** Linearized polynomial \( L(x) = L_1(x) - L_2(x) \) has \( \alpha_1, \ldots, \alpha_d \) as roots. Furthermore, as \( L(x) \) is linear, then also all \( [d] \) combination of \( \alpha_1, \ldots, \alpha_d \) are roots of \( L(x) \). Since the degrees of \( L_1(x) \) and \( L_2(x) \) were smaller than \( [d] \), this holds only if \( L(x) \equiv 0 \). \( \square \)
5.2 A Reed-Solomon like code construction algorithm

Extension field $F = F_q^n$ can be regarded as a vector space of dimension $n$ over $F_q$. Let $A = \{\alpha_1, \alpha_2, \ldots, \alpha_l\} \subset F$ be a set of linearly independent elements of this vector space and $\langle A \rangle \subset F$ be a span of $A$. As an ambient space $W$, we take the direct sum $W = (A) \oplus F = \{ (\alpha, \beta) : \alpha \in \langle A \rangle, \beta \in F \}$, which is a vector space of dimension $l + n$.

Let $F^k[x]$ denote the set of linearized polynomials over $F$ of degree at most $[k - 1]$. If the information vector $U = (u_0, u_1, \ldots, u_{k-1}) \in F^k$ is fixed, then the related function $f(x) \in F^k[x]$ is written as

$$f(x) = \sum_{i=0}^{k-1} u_i x^i.$$

If $A$ is linearly independent set, then also $E = \{(\alpha_i, \beta_i) : \alpha_i \in A, \beta_i = f(\alpha_i)\}$ is linearly independent set with size $l$. Elements of set $E$ span a $l$-dimensional vector space $V = \langle E \rangle \subset W$.

We denote the map which takes $f(x) \in F^k[x]$ to the linear space $V \in \mathcal{P}(W, |A|)$ as $\text{ev}_A$. The vector space produced by map $\text{ev}_A$ on $f(x) \in F^k[x]$ is $\text{ev}_A(f) = \langle \{(\alpha_i, \beta_i) : i \in \{1, \ldots, l\}, \alpha_i \in A, \beta_i = f(\alpha_i)\} \rangle$.

The following two lemmas build a foundation for the final code construction.

**Lemma 4.** If $|A| \geq k$ then the map $\text{ev}_A : F^k[x] \rightarrow \mathcal{P}(W, |A|)$ is injective.

**Proof.** Let $\text{ev}_A(f) = \text{ev}_A(g)$ for $f(x), g(x) \in F^k[x]$. For $h(x) = f(x) - g(x)$, we have $h(\alpha_i) = 0$ for $\alpha_i \in A$ and also $h(x) = 0$ for all $x$ in $\langle A \rangle$. There are $q^{|A|}$ elements in $\langle A \rangle$ and $q^{|A|} \geq q^k$. As the degree of $h(x)$ is at most $k - 1$, then $h(x) \equiv 0$ and $f(x) \equiv g(x)$. \qed

As there are $q^{nk}$ different information vectors $u \in F^k$, then from the Lemma 4 the image of $F^k[x]$ is a code with $q^{nk}$ different codewords.

**Lemma 5.** If $\{(\alpha_1, \beta_1), \ldots, (\alpha_r, \beta_r)\} \subset W$ is a collection of $r$ linearly independent elements satisfying $\beta_i = f(\alpha_i)$ for some linearized polynomial $f$ over $F$, then $\{(\alpha_1, \ldots, \alpha_r)\}$ is linearly independent set.

**Proof.** If $\{\alpha_1, \ldots, \alpha_r\}$ was not linearly independent set, then there would exist nontrivial coefficients $\gamma_1, \ldots, \gamma_r$ such that $\sum_{i=1}^r \gamma_i \alpha_i = 0$. As $f(x)$ is linear, then also $\sum_{i=1}^r \gamma_i (\alpha_i, \beta_i) = 0$. \qed

We saw in the proof of Lemma 4 that the code $C$ from the image under $\text{ev}_A$ of $F^k[x]$ has $q^{nk}$ codewords. As the cardinality of $W$ is $n + l$ and the dimension of $A$ is $l$, then we need only to find the minimum distance of $C$.

**Theorem 8.** Let $C$ be the image under $\text{ev}_A$ of $F^k[x]$ with $l = |A| \geq k$. Then $C$ is a code of type $[n + l, l, nk, 2(l - k + 1)]$. 

14
Proof. We need only to prove the minimum distance as the other parameters result from the discussion before the theorem.

Let \( f(x) \) and \( g(x) \) be distinct elements of \( \mathbb{F}^k[x] \), \( V = \text{ev}_A(f) \) and \( U = \text{ev}_A(g) \). Let the dimension of the intersection \( V \cap U \) be \( r \). It is possible to find \( r \) linearly independent elements \((\alpha_1, \beta_1), \ldots, (\alpha_r, \beta_r)\) such that \( f(\alpha_i) = g(\alpha_i) = \beta_i \) for \( i \in \{1, \ldots, r\} \). From Lemma 5, \( \alpha_1, \ldots, \alpha_r \) are linearly independent and they span an \( r \)-dimensional vector space \( B \) with the property that \( f(b) - g(b) = 0 \) for all \( b \in B \). If \( r \geq k \), then \( f(x) \) and \( g(x) \) are linearized polynomials of degree less than \( q^k \) which agree on at least \( k \) points and thus \( f(x) \equiv g(x) \). This is a contradiction as we assumed that \( f(x) \) and \( g(x) \) are two distinct elements of \( \mathbb{F}^k[x] \), so \( r \leq k - 1 \). From the definition of distance we have

\[
d(V, U) = \dim(V) + \dim(U) - 2 \dim(V \cap U) = 2(l - r) \geq 2(l - k + 1).
\]

The Singleton bound for code with these parameters is

\[
|C| \leq \left[ \frac{n + k}{k} \right]_q < 4q^{kn}.
\]

On the other hand, if we consider the Singleton bound using normalized parameters, we achieve the rate (we denote by \( N = n + l \) and \( D = 2(l-k+1) \))

\[
R = \frac{nk}{Nl} = \frac{kN - kl}{Nl} = \frac{k}{l} - \frac{k}{N} = \frac{k}{l} - \frac{\lambda k}{l} = (1 - \lambda) \left( 1 - \frac{1}{l} + \frac{k}{l} \right) = (1 - \lambda) \left( 1 - \frac{D}{2} \right) = (1 - \lambda) \left( 1 - \delta + \frac{1}{\lambda N} \right).
\]

This rate achieves the Singleton bound in normalized parameters if \( N \to \infty \).

6 Conclusion

In this report, we introduced the concept of random network coding. We showed that vector spaces can be used for transmitting the messages, when interpreted as codewords.

We surveyed bounds on the parameters of such codes and presented a code construction method.

References