Applications of quasi-uniform codes to storage

Thomas, Eldho K.; Oggier, Frédérique

2014

Thomas, E. K. & Oggier, F. (2014). Applications of quasi-uniform codes to storage. 2014 International conference on signal processing and communications (SPCOM), 1-5.

<https://hdl.handle.net/10356/103619>

<https://doi.org/10.1109/SPCOM.2014.6983940>

© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [<http://dx.doi.org/10.1109/SPCOM.2014.6983940>].

Downloaded on 26 Apr 2023 20:40:41 SGT

Applications of Quasi-uniform Codes to Storage

Eldho K. Thomas and Frédérique Oggier
 Division of Mathematical Sciences
 School of Physical and Mathematical Sciences
 Nanyang Technological University, Singapore.
 Email: eldho1@e.ntu.edu.sg, frederique@ntu.edu.sg

Abstract—We consider the design of quasi-uniform codes from dihedral 2-groups. Quasi-uniform codes have the distinctive feature of allowing codeword coefficients to live in different alphabets. We obtain a bound on the minimum distance of quasi-uniform codes coming from p -groups as a function of the number of p -ary codeword components. We study possible applications of such codes to storage, where the minimum distance is important to allow object retrieval, yet binary coefficients are preferred for fast computations, for example during repairs.

I. INTRODUCTION

Traditional erasure (error correction) codes are linear codes defined over finite fields. Motivated by the emergence of new communication technologies (e.g., wireless channels, network coding, distributed storage) or mathematical curiosity, new constructions of codes are constantly appearing, catering to the needs of new communication channels, or providing generalizations of known codes.

In a very general setting, a code may be defined as follows:

Definition 1: A code C of length n is an arbitrary nonempty subset of $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ where $\mathcal{X}_i$ is the alphabet for the i th codeword symbol, and each $\mathcal{X}_i$ might be different.

We are here interested in quasi-uniform codes, whose codeword coefficients indeed live in possibly different alphabets. Quasi-uniform codes are defined with respect to an underlying probability distribution which is quasi-uniform [1], as will be explained in Section II, and quasi-uniform distributions (and in turn quasi-uniform codes [2]) can be constructed from finite groups and their subgroups [3]. When quasi-uniform codes come from finite groups, the underlying group structure can be exploited to derive code properties, such as the minimum distance [8].

In Section III, we present a family of quasi-uniform codes built from the dihedral 2-groups. In this case, codeword coefficients are either binary, or living in an alphabet of size a power of 2. We compute the code parameters, in particular its minimum distance, and study how the behavior of the minimum distance with respect to the coefficient alphabets.

We give concrete code instances in Section IV, where we discuss how such codes can actually be encoded. Properties of quasi-uniform codes coming from dihedral 2-groups include: (1) binary, or power of 2 alphabets, (2) good minimum distance which is easily characterized by the subgroup structure of the group used, (3) control of how many coefficients are binary. These properties motivate us to consider their

applications to storage, which is discussed in Section V, with an explanation of the storage model that we assume.

Finally, a bound on the minimum distance is given for a class of p -groups, which takes into account the number of p -ary coefficients of the codeword.

II. QUASI-UNIFORM CODES FROM GROUPS

Let $\mathcal{A}$ denote a subset of indices from $\mathcal{N} = \{1, \dots, n\}$.

Definition 2: [1] A probability distribution over a set of n jointly distributed discrete random variables $X_1, \dots, X_n$ is said to be *quasi-uniform* if for any $\mathcal{A} \subseteq \mathcal{N}$, $X_{\mathcal{A}} = \{X_i, i \in \mathcal{A}\}$ is uniformly distributed over its support $\lambda(X_{\mathcal{A}}) = \{x_{\mathcal{A}} : P(X_{\mathcal{A}} = x_{\mathcal{A}}) > 0\}$:

$$P(X_{\mathcal{A}} = x_{\mathcal{A}}) = \begin{cases} 1/|\lambda(X_{\mathcal{A}})| & \text{if } x_{\mathcal{A}} \in \lambda(X_{\mathcal{A}}), \\ 0 & \text{otherwise.} \end{cases}$$

We can associate to every code C of length n (see Definition 1) a set of random variables [2] by treating each codeword $(X_1, \dots, X_n) \in C$ as a random vector with probability

$$P(X_{\mathcal{N}} = x_{\mathcal{N}}) = \begin{cases} 1/|C| & \text{if } x_{\mathcal{N}} \in C, \\ 0 & \text{otherwise.} \end{cases}$$

To the i th codeword symbol then corresponds a *codeword symbol random variable* X_i induced by C .

Definition 3: [2] A code C is said to be *quasi-uniform* if the induced codeword symbol random variables are quasi-uniform.

Given a code, we thus know how to associate a set of (quasi-uniform or not) random variables. Conversely, given a set of quasi-uniform random variables $X_1, \dots, X_n$ with probabilities $P(X_{\mathcal{A}} = x_{\mathcal{A}}) = 1/|\lambda(X_{\mathcal{A}})|$ for all $\mathcal{A} \subseteq \mathcal{N}$, the corresponding quasi-uniform code C of length n is given by $C = \lambda(X_{\mathcal{N}}) = \{x_{\mathcal{N}} = P(X_{\mathcal{N}} = x_{\mathcal{N}}) > 0\}$.

Since quasi-uniform distributions may be obtained from finite groups [3], finite groups in turn give rise to quasi-uniform codes [8]. Let us shortly recall these constructions.

Let G be a finite group of order $|G|$ with n subgroups $G_1, \dots, G_n$, and $G_{\mathcal{A}} = \cap_{i \in \mathcal{A}} G_i$. Given a subgroup G_i of G , the (left) coset of G_i in G is defined by $gG_i = \{gh, h \in G_i\}$. The number $[G : G_i]$ of (left) cosets of G_i in G is called the index of G_i in G and it is known from Lagrange Theorem that $[G : G_i] = |G|/|G_i|$. Let X be a random variable uniformly distributed over G , that is $P(X = g) = 1/|G|$, for any $g \in G$. Define the new random variable $X_i = XG_i$, with support the

$[G : G_i]$ cosets of G_i in G . Then $P(X_i = gG_i) = |G_i|/|G|$ and $P(X_i = gG_i, i \in \mathcal{A}) = |\cap_{i \in \mathcal{A}} G_i|/|G|$. More precisely:

Theorem 1: [3] For any finite group G and any subgroups $G_1, \dots, G_n$ of G , there exist n jointly distributed quasi-uniform discrete random variables $X_1, \dots, X_n$ such that for all non-empty subsets $\mathcal{A}$ of $\mathcal{N}$, $P(X_{\mathcal{A}} = x_{\mathcal{A}}) = |G_{\mathcal{A}}|/|G|$.

Quasi-uniform codes are now obtained from these quasi-uniform distributions by taking the support $\lambda(X_{\mathcal{N}})$. Codewords (of length n) can then be described explicitly by letting the random variable X take every possible values in the group G , and by computing the corresponding cosets as follows:

	G_1	$\dots$	G_n
g_1	$g_1 G_1$		$g_1 G_n$
g_2	$g_2 G_1$		$g_2 G_n$
$\vdots$	$\vdots$		$\vdots$
$g_{ G }$	$g_{ G } G_1$	$\dots$	$g_{ G } G_n$

Each row corresponds to one codeword of length n . However, as shown in [8], care should be taken to have $|G_{\mathcal{N}}| = 1$ to ensure that $|C| = |G|$, otherwise $|C| = |G|/|G_{\mathcal{N}}|$ and the above table yields several copies of the same code.

If G_i is normal, the set of cosets $G/G_i := \{gG_i, g \in G\}$ is a group, called quotient group. Thus if G is an abelian group with subgroups $G_1, \dots, G_n$, or if G is non-abelian group, with normal subgroups $G_1, \dots, G_n$, we obtain a quasi-uniform code C which has itself a group structure (with respect to the group structures componentwise). If G is abelian, we furthermore have a labeling of the codewords coefficients which may take values in different abelian groups.

We can easily express the minimum distance of the code in terms of the subgroups $G_1, \dots, G_n$.

Lemma 1: [8] The minimum distance $\min_{c \in C} wt(c)$ of a quasi-uniform code C generated by a (possibly) nonabelian group G and its normal subgroups $G_i, i = 1, \dots, n$ is

$$n - \max_{\mathcal{A} \in \mathcal{N}, G_{\mathcal{A}} \neq \{0\}} |\mathcal{A}|,$$

where the weight $wt(c)$ is understood as the number of components which are not an identity element in the respective group where each coefficient takes values.

III. QUASI-UNIFORM CODES FROM DIHEDRAL GROUPS

Consider the dihedral group $D_{2^{k+1}} = \langle r, s : r^{2^k} = s^2 = 1, rs = sr^{-1} \rangle$ of order 2^{k+1} . A quasi-uniform code C of length n built from $D_{2^{k+1}}$ consists of codewords of length n , where the i th coefficient takes values in the set of cosets of the subgroup $G_i, i = 1, \dots, n$. Since not every subgroup of $D_{2^{k+1}}$ is normal, not every coefficient alphabet will end up having a group structure. However, since we are interested in the subgroup structure, and not in the non-abelian binary operation of this dihedral group, we may alternatively consider its abelian representation [7], that is an abelian group A with subgroups $A_1, \dots, A_n$ such that $|G|/|G_{\mathcal{A}}| = |A|/|A_{\mathcal{A}}|$ for every choice of $\mathcal{A}$ in $\mathcal{N}$. The abelian group representation of

$D_{2^{k+1}}$ [7] is explicitly given by

$$\psi : D_{2^{k+1}} \rightarrow A = (\mathbb{Z}_2)^{k+1} \\ r^a s^j \mapsto \sum_{i=0}^{k-1} a_i e_i + j e_k \quad (1)$$

where a_i 's are obtained from the binary representation of $a = \sum_{i=0}^{k-1} a_i 2^i$ and $(e_0, \dots, e_k)$ is the standard basis of $(\mathbb{Z}_2)^{k+1}$. The code $\psi(C)$ thus has an abelian group structure, but the length n and minimum distance (where the weight is understood as the number of coefficients not equal to G_i) of both codes are the same, since ψ preserves the group structure.

Code Length. To determine the largest value of n , we only need to count the number of non-trivial subgroups of $D_{2^{k+1}}$ (or $\psi(D_{2^{k+1}})$, whichever one finds most convenient). Subgroups of $D_{2^{k+1}}$ are of the form $\langle r^{2^i}, r^c s \rangle$ where $1 \leq i \leq k, 0 \leq c < 2^i$ and $\langle r^{2^{i-1}} \rangle$; $1 \leq i \leq k$ [7]. Note that $|\langle r^{2^i}, r^c s \rangle| = |\langle r^{2^{i-1}} \rangle| = 2^{k-i+1}$, hence $[D_{2^{k+1}} : \langle r^{2^i}, r^c s \rangle] = [D_{2^{k+1}} : \langle r^{2^{i-1}} \rangle] = 2^i$ which implies that there are $2^i + 1$ subgroups of index 2^i . The total number of proper subgroups is then

$$\sum_{i=1}^k 2^i + k = 2[2^k - 1] + k = 2^{k+1} + k - 2.$$

The center of $D_{2^{k+1}}$ is $\langle r^{2^{k-1}} \rangle$, which is the intersection of all subgroups except subgroups of order two (or of index 2^k), thus at least one subgroup of order 2 should be taken to be build a code, to ensure that $G_{\mathcal{N}} = 1$.

Minimum Distance. Consider the quasi-uniform code C of maximum length formed by all non-trivial subgroups of $G = D_{2^{k+1}}$. Its minimum distance is by Lemma 1

$$d = n - \max_{G_{\mathcal{A}} \neq 1} |\mathcal{A}|.$$

Since there are 2^k subgroups of order 2, and that they are the only ones which do not intersect the center, the number of subgroups intersecting the center is $(2^{k+1} + k - 2) - 2^k = 2^k + k - 2$. The minimum distance d is then

$$d = (2^{k+1} + k - 2) - (2^k + k - 2) = 2^k.$$

To summarize, the code parameters of C are

$$(n, |C|, d) = (2^{k+1} + k - 2, 2^{k+1}, 2^k), \quad (2)$$

with alphabet $(\mathbb{Z}_2)^k$ for $2^k + 1$ codeword coefficients.

Remark 1: Quasi-dihedral 2-groups and generalized quaternion groups of order 2^{k+1} also generate abelian representable codes coming from $A = (\mathbb{Z}_2)^{k+1}$ and its subgroups [7]. These two families of 2-groups thus do not provide new code constructions.

IV. QUASI-UNIFORM CODES FROM D_8

Consider the dihedral group $D_8 = \langle r, s : r^4 = s^2 = 1, rs = sr^{-1} \rangle$, whose subgroup lattice diagram is shown in Figure 1.

Now the map $\psi : D_8 \rightarrow \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ from (1) explicitly gives the abelian representation of D_8 (to simplify the notation, we write abc for $(a, b, c) \in \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$):

$$1 \mapsto 000, r \mapsto 100, r^2 \mapsto 010, r^3 \mapsto 110,$$

Fig. 1. The dihedral group D_8 and its lattice of subgroups.

$$s \mapsto 001, rs \mapsto 101, r^2s \mapsto 011, r^3s \mapsto 111,$$

and the subgroups are mapped accordingly:

$$\begin{aligned} \langle r \rangle &\mapsto \langle 100, 010 \rangle = G_1, & \langle r^2, s \rangle &\mapsto \langle 010, 001 \rangle = G_2, \\ \langle r^2, rs \rangle &\mapsto \langle 010, 101 \rangle = G_3, & \langle r^3s \rangle &\mapsto \langle 111 \rangle = G_4, \\ \langle s \rangle &\mapsto \langle 001 \rangle = G_5, & \langle r^2 \rangle &\mapsto \langle 010 \rangle = G_6, \\ \langle r^2s \rangle &\mapsto \langle 011 \rangle = G_7, & \langle rs \rangle &\mapsto \langle 101 \rangle = G_8, \end{aligned}$$

resulting in a lattice of subgroups of $\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ which is identical to that of D_8 . We compute next the corresponding code.

Code Construction. Since G_1, G_2, G_3 have order 4, their corresponding quotient groups have order 2 and are isomorphic to $\mathbb{Z}_2$. The other proper subgroups have order 2, and thus the corresponding quotient groups are $\mathbb{Z}_2 \oplus \mathbb{Z}_2$.

Pick for example $g = 100 \in \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$. Then $g + G_1 = 100 + \langle 100, 010 \rangle = 100 + \{100, 010, 000, 110\} = G_1$. This is thus the identity element of the quotient group. Next $g + G_6 = 100 + \langle 010 \rangle = 100 + \{010, 000\}$, which we map to 10 (the Klein group $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ has 3 elements of order 2, thus any choice will do). What matters for the minimum distance is the identity element, for which there is no choice.

The code is shown in Table I, where we have used that $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ contains an isomorphic copy of $\mathbb{Z}_2$, namely $\mathbb{Z}_2 \oplus \{0\}$, to emphasize binary coefficients. From (2), which can be seen directly as well, its parameters are a length of $n = 8$, a cardinality $|C| = 8$ and a minimum distance of $d = 4$.

Encoding. The term “generator matrix” is probably risky in this context, since we do not have a vector space, however, we do have a $\mathbb{Z}_2$ -basis for the code, that is, a set of linearly independent vectors such that any binary linear combination generates the code. By defining $a(b, c) = (ab, ac)$ for $a, b, c \in \mathbb{Z}_2$, we may encode the above codewords by

$$[u_1, u_2, u_3] \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 01 & 01 & 01 & 01 & 0 & 0 & 0 \\ 01 & 0 & 01 & 1 & 11 & 1 & 0 & 1 \end{bmatrix} \quad (3)$$

for $u_1, u_2, u_3 \in \mathbb{Z}_2$.

$\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$	G_6	G_5	G_7	G_8	G_4	G_1	G_2	G_3
000	0	0	0	0	0	0	0	0
100	1	1	1	1	1	0	1	1
010	0	01	01	01	01	0	0	0
110	1	11	11	11	11	0	1	1
001	01	0	01	1	11	1	0	1
101	11	1	11	0	01	1	1	0
011	01	01	0	11	1	1	0	1
111	11	11	1	01	0	1	1	0

TABLE I
A $(8, |C|, 3)$ CODE CONSTRUCTED FROM D_8 , $|C| = 8$. PAIRS ARE ELEMENTS IN $\mathbb{Z}_2 \oplus \mathbb{Z}_2$.

V. STORAGE APPLICATIONS

Suppose we have n nodes across which data is stored using an erasure code C of length n . Then every coefficient of a codeword of C is stored at every node. In case of node failure(s), we should be able to retrieve the stored data. Whether this is possible depends on the number of failures and on the minimum distance of the code. We thus start by comparing the minimum distance of the proposed codes to known codes.

A. Code Comparisons

We compare in Table II the codes obtained from dihedral groups with binary codes, more precisely, if a dihedral code has length n and cardinality 2^{k+1} , we compare it with a binary $(n, k+1, d)$ code, where $k+1$ is the dimension of the code.

- When $k = 2$, the code from D_8 has the same minimum distance as a binary code. When this code is punctured at the components corresponding to G_1, G_2, G_3 respectively, the minimum distance stays $d = 4$, but so is the minimum distance of the binary code when punctured.
- When $k = 3$, the code from D_{16} also has the same minimum as a binary code. However the behavior improves after puncturing at least three of the components corresponding to subgroups of order 8.

Since the finite field $\mathbb{F}_4$ can be seen as a vector space over $\mathbb{F}_2$ by fixing a $\mathbb{F}_2$ -basis (we write $\mathbb{F}_2$ to emphasize the field structure, and $\mathbb{Z}_2$ to emphasize the additive group one), we could see the coefficients of the code in $\mathbb{F}_4$. We also put parameters of codes over $\mathbb{F}_4$ in Table II, however a code over $\mathbb{F}_4$ contains 4^{k+1} , so for a fair comparison, we should really compute the minimum distance of the proposed codes as codes over $\mathbb{F}_4$ (which we skip here, since we are interested in the minimum distance as predicted by the subgroup structure).

B. A Storage Example

We illustrate how the codeword (3) can be used over $n = 8$ nodes. Suppose a data object (u_1, u_2, u_3) needs to be stored. Consider the following storage allocation:

$(n, k+1)$	d over $\mathbb{F}_2$	d over $\mathbb{F}_4$ [5]	New d
(8, 3)	4	5	4
(7, 3)	4	4	4
(6, 3)	4	4	4
(5, 3)	4	3	4
(17, 4)	8	12	8
(16, 4)	8	11	8
(15, 4)	8	10	8
(14, 4)	7	9	8
(13, 4)	6	8	8
(12, 4)	6	7	8
(11, 4)	5	6	8
(10, 4)	4	6	8
(9, 4)	4	5	8

TABLE II
MINIMUM DISTANCE COMPARISON WITH KNOWN CODES [5].

node 1: (u_1, u_3) node 5: $(u_1 + u_3, u_2 + u_3)$
node 2: (u_1, u_2) node 6: $(u_3, 0)$
node 3: $(u_1, u_2 + u_3)$ node 7: $(u_1, 0)$
node 4: $(u_1 + u_3, u_2)$ node 8: $(u_1 + u_3, 0)$

We easily notice that three failures at most can be tolerated (corresponding to a minimum distance of 4 indeed): u_2 is present only in nodes 2,3,4 and 5, so if these four nodes fail, u_2 is lost.

In case of one node failure, this node is repaired easily: indeed, this codeword is created from $\mathbb{Z}_2$ -linear combinations. For example:

- node 1 is repaired by downloading u_1 (from node 2 or node 7) and u_3 (from node 6),
- node 2 is repaired by downloading u_2 from node 4 and u_1 (from node 1 or 7).

Apart repairs obtained using $\mathbb{Z}_2$ -linear combination, a distinctive feature of this code is the fact that we control how many coefficients live in $\mathbb{Z}_2$, and how many live in $\mathbb{Z}_2 \oplus \mathbb{Z}_2$.

VI. BOUNDS ON THE MINIMUM DISTANCE

It is known that the minimum distance of a linear code over $\mathbb{F}_q$ depends on the choice of q . The same phenomenon appears here: the minimum distance depends on the different alphabets in which live the codeword coefficients. In particular, the more coefficients we want small, the more subgroups of high index we need, which reduces the minimum distance. In this section, we will give a bound on the minimum distance of a quasi-uniform code, when it is obtained from a p -group as described above from a dihedral 2-group (that is using its abelian group representation A).

Definition 4: The Frattini subgroup $\phi(G)$ of a group G is the intersection of all maximal subgroups of G . In particular, $\phi(G) = G$, if there are no maximal subgroups. It is also known that $\phi(G)$ is normal in G . [4]

To prove the main result of this section, we need the following lemmas.

Lemma 2: [6] Every maximal subgroup of a finite p -group is normal with index p .

Lemma 3: Let $\phi(G)$ be the Frattini subgroup of a finite p -group G . Then $G/\phi(G)$ is elementary abelian.

Proof: If M is maximal in G , then G/M is of order p by Lemma 2, and G/M is abelian. The commutator subgroup G' is then a subgroup of M for all maximal subgroups $M \leq G$ and then $G' \leq \phi(G)$ implies that $G/\phi(G)$ is abelian [4].

Since $|G/M| = p$, $(gM)^p = M$ for all $g \in G$, that is, $g^p M = M$ for all $g \in G$ and for all maximal subgroups M . Hence $g^p \in G/\phi(G)$ and $g^p \phi(G) = \phi(G)$. Therefore, if $g\phi(G) \in G/\phi(G)$, its order $|g\phi(G)|$ is p for all $g \in G$, and $G/\phi(G)$ is an abelian group where each non-trivial element has order p . ■

The next lemma counts the subgroups of order p of an elementary abelian p -group.

Lemma 4: Let G be an elementary abelian p -group of order p^m . Then there are $\frac{p^m - 1}{p - 1}$ subgroups of order p .

Proof: An elementary abelian group is of the form $(\mathbb{Z}_p)^m$, by the classification of finitely generated abelian groups [4]. Let G_i be a subgroup of order p . Since p is a prime, G_i is cyclic, spanned by a single element in G .

Now each element in G has m components and each component takes p values and therefore the total number of subgroups spanned by elements of the form $\langle(0, \dots, i, 0, \dots, 0)\rangle$ is m .

Similarly, the total number of subgroups spanned by elements of the form $\langle(0, \dots, 1, 0, \dots, i, 0, \dots, 0)\rangle$ is $\binom{m}{2}(p - 1)$ since i varies from 0 to $p - 1$. (Also note that $\langle(0, \dots, i, 0, \dots, 1, 0, \dots, 0)\rangle = \langle(0, \dots, 1, 0, \dots, j, 0, \dots, 0)\rangle$.)

Using similar type of arguments the total number of subgroups of order p is $n = m + \binom{m}{2}(p - 1) + \binom{m}{3}(p - 1)^2 + \dots + \binom{m}{m-1}(p - 1)^{m-2} + \binom{m}{m}(p - 1)^{m-1}$. After simplification, $n = \frac{1}{p-1} [\sum_{i=0}^m \binom{m}{i} (p-1)^i - 1] = \frac{1}{p-1} [(1 + (p-1))^m - 1] = \frac{p^m - 1}{p - 1}$. ■

From Lemma 3 and Lemma 4, it is clear that there are $\frac{p^m - 1}{p - 1}$ subgroups of $G/\phi(G)$ of order p .

Proposition 1: The number of maximal subgroups of a p -group having Frattini subgroup of index p^m is $\frac{p^m - 1}{p - 1}$.

Proof: Let G be a group of order p^k and suppose $[G : \phi(G)] = p^m$. By Lemma 4, the number of subgroups of order p of the quotient group $G/\phi(G)$ is $\frac{p^m - 1}{p - 1}$. Now by the subgroup correspondence theorem [4], the subgroups of G containing $\phi(G)$ and subgroups of $G/\phi(G)$ are in 1-1 correspondence.

Let M be a maximal subgroup of G of index p (Lemma 2). Then $|M| = p^{k-1}$. Consider the subgroup, $M/\phi(G) \leq G/\phi(G)$. Since $|G/\phi(G)| = p^m$ and $|M/\phi(G)| = p^{k-1-(k-m)} = p^{m-1}$ and therefore $[G/\phi(G) : M/\phi(G)] = p$. From [4] and [9], the number of index p -subgroups and the number of order p -subgroups of an elementary abelian p -group are the same, $\frac{p^m - 1}{p - 1}$ in total. ■

Let G be a p -group which is abelian representable. Recall that it means that we can find an abelian group A and subgroups A_i such that $[G : G_{\mathcal{A}}] = [A : A_{\mathcal{A}}]$ for all $\mathcal{A} \subseteq \mathcal{N}$. Note that A may have more subgroups than G .

Suppose that the Frattini subgroup $\phi(G)$ of G is non-trivial and of index p^m . Because A may contain more subgroups than G , $|\phi(G)|$ and $|\phi(A)|$ may not be equal. For example, $|\phi(D_8)| = 2 \neq 1 = |\phi(\mathbb{Z}_2^3)|$. By Proposition 1, there are $\frac{p^m-1}{p-1}$ maximal subgroups of G of index p . Suppose $\psi : G \rightarrow A$ is a mapping which preserves the subgroup structure of G , used to prove that A is an abelian group representation of G . Under this mapping, the subgroup lattice of G is embedded into that of A and $[G : \phi(G)] = [A : \psi(\phi(G))] = p^m$.

Suppose we construct a quasi-uniform code C of length n from G using its abelian representation A and assume that all subgroups corresponding to the maximal subgroups of G are included. Then the minimum distance of C is bounded as follows:

Proposition 2: Let G be a p -group which is abelian representable and suppose G has a non-trivial Frattini subgroup, $\phi(G)$ of index p^m . Then the quasi-uniform code obtained from the abelian representation of G has a minimum distance d such that:

$$1 \leq d \leq n - \frac{p^m - 1}{p - 1},$$

where n is the length of the code.

Proof: Since no codeword is repeated, $G_N = 1$ implies $n > \max_{G_A \neq 1} |\mathcal{A}|$ so that $d = n - \max_{G_A \neq 1} |\mathcal{A}| \geq 1$.

From Proposition 1, $\max_{G_A \neq 1} |\mathcal{A}| \geq \frac{p^m-1}{p-1}$ since all maximal subgroups are included and hence $d \leq n - \frac{p^m-1}{p-1}$. That is,

$$1 \leq d \leq n - \frac{p^m - 1}{p - 1}.$$

The above bound is tight. Consider the $(7, |C|, 4)$ code C , obtained by puncturing the first column of the $(8, |C|, 4)$ code proposed in Section IV. That is, the puncturing is done by removing a subgroup which is not maximal. Its minimum distance is $d = 4$. Since $[D_8 : \phi(D_8)] = 4$ and $n - (2^2 - 1) = n - 3 = 4 = d$, equality holds.

Corollary 1: Consider the quasi-uniform code construction of the above proposition, but where we do not include l maximal subgroups out of $\frac{p^m-1}{p-1}$ in the code construction. Then $\max_{G_A \neq 1} |\mathcal{A}| \geq \frac{p^m-1}{p-1} - l$ and the bound on the minimum distance becomes:

$$1 \leq d \leq n + l - \frac{p^m - 1}{p - 1}.$$

It is worth repeating that the number of maximal subgroups also decide the number of components of the codewords that will live in $\mathbb{Z}_p$, thus this bound gives a trade-off between coefficients in $\mathbb{Z}_p$ and minimum distance.

Corollary 2: Let C be a quasi-uniform code of length n formed from the dihedral 2-group $D_{2^{k+1}}$. Suppose t subgroups of order 2, distinct from the center, are included. Then the minimum distance d of the code is $d = t \leq 2^k$, that is, it only depends on the number of subgroups of order 2.

Proof: We know that the center $\langle r^{2^{k-1}} \rangle$ is of order 2 and is contained in all but subgroups of order 2. Then $\max_{G_A \neq 1} |\mathcal{A}| = n - t$ implies $d = n - (n - t) = t$. There are

2^k subgroups of order two except the center and so $t \leq 2^k$. ■

ACKNOWLEDGEMENT

The work is supported by the Nanyang Technological University under Research Grant M58110049.

REFERENCES

- [1] T.H. Chan, "A Combinatorial Approach to Information Inequalities", *Commun. Inf. Syst.*, vol. 1, 2001.
- [2] T.H. Chan, A. Grant and T. Britz, "Properties of Quasi-Uniform Codes", *2010 IEEE International Symposium on Information Theory (ISIT 2010)*, Austin, TX, 13-18 June 2010.
- [3] T.H. Chan and R. W. Young, "On a Relation Between Information Inequalities and Group Theory", *IEEE Trans. Inform. Theory*, vol. 48, pp. 1992-1995, 2002.
- [4] D. Dummit and R. Foote, "Abstract Algebra", John Wiley and Sons, 2004.
- [5] M. Grassl, "Bounds on the minimum distance of linear codes and quantum codes", available at <http://www.codetables.de>, Accessed on 2014-04-24.
- [6] K. Igusa, "p-groups", Lecture notes, available at <http://people.brandeis.edu/~igusa/Math101b/pGroups.pdf>.
- [7] E. Thomas, N. Markin and F. Oggier, "On Abelian Group Representability of Finite Groups", *Advances in Mathematics of Communications*, 2013, available at <http://arxiv.org/abs/1212.1298>.
- [8] E. Thomas, F. Oggier, "Explicit Constructions of Quasi-Uniform Codes from Groups", *2013 IEEE International Symposium on Information Theory (ISIT 2013)*, Istanbul.
- [9] <http://crazyproject.wordpress.com/2010/06/16/properties-of-the-image-and-kernel-of-the-p-power-map-on-a-finite-abelian-group/>.