Thirteenth International Workshop on Algebraic and Combinatorial Coding Theory
June 15-21, 2012, Pomorie, Bulgaria *pp. 329–334*

Bounds on list decoding Gabidulin codes

Antonia Wachter-Zeh[1] antonia.wachter@uni-ulm.de
Institute of Communications Engineering, Ulm University, Ulm, Germany and Institut de Recherche Mathématique de Rennes, Université de Rennes 1, Rennes, France

Abstract. An open question about Gabidulin codes is whether polynomial-time list decoding beyond half the minimum distance is possible or not. In this contribution, we give a lower and an upper bound on the list size, i.e., the number of codewords in a ball around the received word. The lower bound shows that if the radius of this ball is greater than the Johnson radius, this list size can be exponential and hence, no polynomial-time list decoding is possible. The upper bound on the list size uses subspace properties.

1 Introduction

Gabidulin codes [1] can be seen as the analogs of Reed–Solomon (RS) codes in rank metric. There are several efficient decoding algorithms up to half the minimum rank distance. However, in contrast to RS codes, there is no polynomial-time decoding algorithm beyond half the minimum distance. For RS codes, it can be shown that the number of codewords in a ball around *any* received word is always polynomial in the length when the radius of the ball is at most the Johnson radius. The Guruswami–Sudan algorithm provides an efficient polynomial-time list decoding algorithm of RS codes up to the Johnson radius.

For Gabidulin codes, there is no polynomial-time list decoding algorithm; it is not even known, whether such an algorithm can exist or not. An exponential lower bound on the number of codewords in a ball of radius τ around the received word would prohibit polynomial-time list decoding since the list size can be exponential, whereas a polynomial upper bound would show that it might be possible. Faure [2] and Augot and Loidreau [3] made first investigations of this problem.

In this paper, we provide a lower and an upper bound on the list size. The lower bound shows that the list size can be exponential in the length when the radius is at least the Johnson radius and therefore in this region, no polynomial-time list decoding is possible. The upper bound uses the properties of subspaces and gives a good estimate of the number of codewords in such a ball, but is exponential in the length and therefore does not provide an answer to polynomial-time list decodability in the region up to the Johnson bound.

[1] This work was supported by the German Research Council "Deutsche Forschungsgemeinschaft" (DFG) under Grant No. Bo 867/21-2.

2 Preliminaries

2.1 Definitions and notations

Let q be a power of a prime, let $\mathbb{F}_q$ denote the finite field of order q and let $\mathbb{F}_{q^m}$ be the extension field of degree m over $\mathbb{F}_q$. We denote $x^{[i]} = x^{q^i}$ for any integer i, then a *linearized polynomial*, introduced by Ore [4], over $\mathbb{F}_{q^m}$ has the form $f(x) = \sum_{i=0}^{d_f} f_i x^{[i]}$, with $f_i \in \mathbb{F}_{q^m}$. If the coefficient $f_{d_f} \neq 0$, we call $d_f \overset{\text{def}}{=} \deg_q f(x)$ the *q-degree* of $f(x)$. For all $\alpha_1, \alpha_2 \in \mathbb{F}_q$ and $\forall\ a, b \in \mathbb{F}_{q^m}$, the following holds: $f(\alpha_1 a + \alpha_2 b) = \alpha_1 f(a) + \alpha_2 f(b)$. The (usual) addition and the non-commutative composition $f(g(x))$ (also called *symbolic product*) convert the set of linearized polynomials into a non-commutative ring with the identity element $x^{[0]} = x$. In the following, all polynomials are linearized polynomials.

Given a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, there exists a one-to-one mapping for each vector $\mathbf{x} \in \mathbb{F}_{q^m}^n$ on a matrix $\mathbf{X} \in \mathbb{F}_q^{m\times n}$. Let $\operatorname{rank}(\mathbf{x})$ denote the (usual) rank of $\mathbf{X}$ over $\mathbb{F}_q$ and let $\mathcal{R}(\mathbf{x}) = \mathcal{R}(\mathbf{X})$ denote the row space of $\mathbf{X}$ over $\mathbb{F}_q$. The kernel of a matrix is denoted by $\ker(\mathbf{x}) = \ker(\mathbf{X})$ and the image by $\operatorname{im}(\mathbf{x}) = \operatorname{im}(\mathbf{X})$. For an $m \times n$ matrix, if $\dim\ker(\mathbf{x}) = t$, then $\dim\operatorname{im}(\mathbf{x}) = \operatorname{rank}(\mathbf{x}) = n - t$. Throughout this paper, we use the notation as vector or matrix equivalently, whatever is more convenient. The *minimum rank distance* d of a code $\mathcal{C}$ is defined by

$$d = \min\{\operatorname{rank}(\mathbf{c}) \mid \mathbf{c} \in \mathcal{C}, \mathbf{c} \neq \mathbf{0}\}.$$

A *Gabidulin code* can be defined by the evaluation of degree-restricted linearized polynomials as follows.

Definition 1 (Gabidulin Code [1])**.** *A linear $\mathcal{G}(n,k)$ Gabidulin code of length n and dimension k over $\mathbb{F}_{q^m}$ for $n \leq m$ is the set of all codewords, which are the evaluation of a q-degree restricted linearized polynomial $f(x)$:*

$$\mathcal{G}(n,k) \overset{\text{def}}{=} \{\mathbf{c} = (f(\alpha_0), f(\alpha_1), \dots, f(\alpha_{n-1}) \big| \deg_q f(x) < k)\},$$

where the fixed elements $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{F}_{q^m}$ are linearly independent over $\mathbb{F}_q$.

Gabidulin codes are *Maximum Rank Distance* (MRD) codes, i.e., they fulfill the rank metric Singleton bound with equality and $d = n - k + 1$ [1].

The number of s-dimensional subspaces of an n-dimensional vector space over $\mathbb{F}_q$ is the Gaussian binomial, calculated by

$$\begin{bmatrix} n \\ s \end{bmatrix} \overset{\text{def}}{=} \prod_{i=0}^{s-1} \frac{q^n - q^i}{q^s - q^i},$$

with the upper and lower bounds $q^{s(n-s)} \leq \left[\begin{smallmatrix} n \\ s \end{smallmatrix}\right] \leq 4q^{s(n-s)}$.

Moreover, $\mathcal{B}_\tau(\mathbf{a})$ denotes a ball of radius τ in rank metric around a word $\mathbf{a} \in \mathbb{F}_{q^m}^n$. The volume of $\mathcal{B}_\tau(\mathbf{a})$ is independent of its center and is simply the number of $m \times n$ matrices of rank less than or equal to τ.

2.2 Problem statement

Problem 1 (Maximum List Size)**.** *Let the Gabidulin code $\mathcal{G}(n,k)$ over $\mathbb{F}_{q^m}$ with $n \leq m$ and $d = n-k+1$ be given. Let $\tau < d$. Find a lower and upper bound on the maximum number of codewords ℓ in the ball of rank radius τ around $\mathbf{r} = (r_0 \ r_1 \ \dots \ r_{n-1}) \in \mathbb{F}_{q^m}^n$. Hence, find a bound on*

$$\ell \overset{\text{def}}{=} \max_{\mathbf{r} \in \mathbb{F}_{q^m}^n} \left(|\mathcal{B}_\tau(\mathbf{r}) \cap \mathcal{G}| \right).$$

For an upper bound, we have to show that the bound holds for *any* received word $\mathbf{r}$, whereas for a lower bound it is sufficient to show that there exists one $\mathbf{r}$ for which this bound on the list size is valid.

Let $\mathcal{L} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_\ell\}$ with $|\mathcal{L}| = \ell$ denote the list of all codewords in the ball of radius τ around $\mathbf{r}$, i.e., $\mathbf{c}_i \in \mathcal{G}(n,k)$ and $\operatorname{rank}(\mathbf{r} - \mathbf{c}_i) \leq \tau$, for $i = 1, \dots, \ell$.

3 A lower bound on the list size

Faure showed with a probabilistic approach in [2] that the maximum list size of Gabidulin codes with $n = m$ is exponential in n for $\tau \geq n - \sqrt{n(n-d)}$. Our bound slightly improves this value and uses a different proving strategy, based on evaluating linearized polynomials. This approach is inspired by Justesen and Hoholdt's [5] and Ben-Sasson, Kopparty, and Radhakrishna's [6] approaches for bounding the list size of RS codes.

Theorem 1 (Lower Bound on the List Size)**.** *Let the Gabidulin code $\mathcal{G}(n,k)$ over $\mathbb{F}_{q^m}$ with $n \leq m$ and $d = n-k+1$ be given. Let $\tau < d$. Then, there exists a word $\mathbf{r} \in \mathbb{F}_{q^m}^n$ such that*

$$\ell \geq |\mathcal{B}_\tau(\mathbf{r}) \cap \mathcal{G}| \geq \frac{\begin{bmatrix} n \\ n-\tau \end{bmatrix}}{(q^m)^{n-\tau-k}} = q^m q^{\tau(m+n)-\tau^2-md}, \tag{1}$$

and for the special case of $n = m$: $\ell \geq q^n q^{2n\tau-\tau^2-nd}$.

Proof. Since $\tau < d = n-k+1$, also $k-1 < n-\tau$ holds. Consider all monic linearized polynomials of q-degree exactly $n-\tau$ with a root space of dimension $n-\tau$, where all roots are in $\mathbb{F}_{q^n}$. There are exactly (see e.g. [7, Theorem 11.52]) $\begin{bmatrix} n \\ n-\tau \end{bmatrix}$ such polynomials. Now, let us consider a subset of these polynomials, denoted by $\mathcal{P}$: all polynomials where the q-monomials of q-degree greater than or equal to k have the same coefficients. Due to Dirichlet's principle there exist coefficients such that the number of such polynomials is

$$|\mathcal{P}| \geq \frac{\begin{bmatrix} n \\ n-\tau \end{bmatrix}}{(q^m)^{n-\tau-k}},$$

since there are $(q^m)^{n-\tau-k}$ possibilities to choose the highest $n-\tau-(k-1)$ coefficients of a *monic* linearized polynomial over $\mathbb{F}_{q^m}$.

Note that the difference between any two polynomials in $\mathcal{P}$ is a linearized polynomial of q-degree strictly less than k and therefore the evaluation polynomial of a codeword of $\mathcal{G}(n,k)$. Let $\mathbf{r}$ be the evaluation of $f(x) \in \mathcal{P}$ at a basis $\mathcal{A} = \{\alpha_0, \alpha_1, \dots, \alpha_{n-1}\}$ of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$:

$$\mathbf{r} = (r_0 \; r_1 \; \dots \; r_{n-1}) = (f(\alpha_0) \; f(\alpha_1) \; \dots \; f(\alpha_{n-1})).$$

Further, let also $g(x) \in \mathcal{P}$, then $f(x) - g(x)$ has q-degree less than k. Let $\mathbf{c}$ denote the evaluation of $f(x) - g(x)$ at $\mathcal{A}$. Then, $\mathbf{r} - \mathbf{c}$ is the evaluation of $f(x) - f(x) + g(x) = g(x) \in \mathcal{P}$, whose root space has dimension $n-\tau$ and all roots are in $\mathbb{F}_{q^n}$. Thus, $\dim\ker(\mathbf{r}-\mathbf{c}) = n-\tau$ and $\dim\operatorname{im}(\mathbf{r}-\mathbf{c}) = \operatorname{rk}(\mathbf{r}-\mathbf{c}) = \tau$. Therefore, for *any* $g(x) \in \mathcal{P}$, the evaluation of $f(x) - g(x)$ is a codeword from $\mathcal{G}(n,k)$ and has rank distance τ from $\mathbf{r}$. This provides the following lower bound on the maximum list size:

$$\ell \geq \frac{\left[{n \atop n-\tau}\right]}{(q^m)^{n-\tau-k}} \geq \frac{q^{(n-\tau)\tau}}{(q^m)^{n-\tau-k}} = q^m q^{\tau(m+n)-\tau^2-md},$$

and for $n = m$ the special case follows. □

This lower bound is valid for any $\tau < d$, but we want to know, which is the smallest value for τ such that this expression is *exponential* in n. For $n = m$, we can rewrite (1) by

$$\ell \geq q^{n(1-\epsilon)} \cdot q^{2n\tau-\tau^2-nd+n\epsilon},$$

where the first part is exponential in n for any $0 \leq \epsilon < 1$. The second exponent is positive for

$$\tau \geq n - \sqrt{n(n-d+\epsilon)} \overset{\text{def}}{=} \tau_{LB}. \qquad (2)$$

Therefore, our lower bound (1) shows that the maximum list size is exponential in n for $\tau \geq \tau_{LB}$. For $\epsilon = 0$, the value τ_{LB} gives exactly the Johnson radius for Hamming metric.

This reveals a difference between the known limits to list decoding of Gabidulin and RS codes. For RS codes, polynomial-time list decoding up to the Johnson radius is guaranteed by the Guruswami–Sudan algorithm. However, it is not proven that the Johnson radius is tight for RS codes, i.e., it is not known if the list size is polynomial between the Johnson radius and the known exponential lower bounds (see e.g. [5, 6]). For Gabidulin codes, we have shown that the maximum list size is exponential for $\tau \geq \tau_{LB}$, which is asymptotically equal to the Hamming metric Johnson radius.

Example 1. *For the Gabidulin code $\mathcal{G}(n=12, k=6)$ with $d=7$, the Bounded Minimum Distance decoding radius is $\tau_{BMD} = \lfloor (d-1)/2 \rfloor = 3$, the lower bound* *by Faure (equivalent to the Hamming metric Johnson radius) is* $\tau_J = \lceil 4.2 \rceil = 5$ *and* (2) *with* $\epsilon = 0.9$ *gives* $\tau_{LB} = \lceil 3.58 \rceil = 4$*. This means for this code of rate* $k/n = 1/2$*, no polynomial time list-decoding beyond* τ_{BMD} *is possible.*

4 An upper bound on the list size

The following lemma shows that the row spaces of $\mathbf{r} - \mathbf{c}_i$ and $\mathbf{r} - \mathbf{c}_j$, $\mathbf{c}_i, \mathbf{c}_j \in \mathcal{L}$, $i \neq j$, have no $(2\tau - d + 1)$-dimensional subspace in common.

Lemma 1. *Let* $\tau < d$ *and let* $\mathbf{r} \in \mathbb{F}_{q^m}^n$*. Let* $\mathbf{c}_i$*, for* $i = 1, \dots, \ell$*, be codewords of the Gabidulin code* $\mathcal{G}(n,k)$ *with minimum rank distance* d *and let* $\mathrm{rk}(\mathbf{r} - \mathbf{c}_i) \leq \tau$ *hold for all* $i = 1, \dots, \ell$*. Let* $\mathrm{rk}(\mathbf{r} - \mathbf{c}_i) = t_i \leq \tau$ *and* $\mathrm{rk}(\mathbf{r} - \mathbf{c}_j) = t_j \leq \tau$*,* $i \neq j$*. Then, the row spaces of* $(\mathbf{r} - \mathbf{c}_i)$ *and* $(\mathbf{r} - \mathbf{c}_j)$ *have no subspace of dimension at least* $t_i + t_j - d + 1$ *in common, for* $\lfloor (d-1)/2 \rfloor < t_i, t_j \leq \tau$*.*

Proof. Assume, there exist $(\mathbf{r} - \mathbf{c}_i)$ and $(\mathbf{r} - \mathbf{c}_j)$, with $\mathrm{rank}(\mathbf{r} - \mathbf{c}_i) = t_i$, $\mathrm{rank}(\mathbf{r} - \mathbf{c}_j) = t_j$, $i \neq j$, such that their row spaces contain the same subspace of dimension at least $(t_i + t_j - d + 1)$. Then,

$$\dim(\mathcal{R}(\mathbf{c}_i - \mathbf{c}_j)) = \dim(\mathcal{R}(\mathbf{r} - \mathbf{c}_i - \mathbf{r} + \mathbf{c}_j)) \leq \dim\left(\mathcal{R}\begin{pmatrix}\mathbf{r} - \mathbf{c}_i \\ \mathbf{r} - \mathbf{c}_j\end{pmatrix}\right)$$
$$\leq t_i + t_j - (t_i + t_j - d + 1) = d - 1,$$

which is a contradiction to $\mathrm{rk}(\mathbf{c}_i - \mathbf{c}_j) \geq d$. □

This means in particular, if $\mathrm{rk}(\mathbf{r} - \mathbf{c}_i) = \mathrm{rk}(\mathbf{r} - \mathbf{c}_j) = t \leq \tau$, they have no subspace of dimension at least $2t - d + 1$ in common. Based on this lemma, we obtain the following upper bound on the maximum list size.

Theorem 2 (Upper Bound on the List Size)**.** *Let the Gabidulin code* $\mathcal{G}(n,k)$ *over* $\mathbb{F}_{q^m}$ *with* $n \leq m$ *and* $d = n - k + 1$ *be given. Let* $\tau < d$*. Then, for any word* $\mathbf{r} \in \mathbb{F}_{q^m}^n$ *and hence, for the maximum list size, the following holds*

$$\ell = \max_{\mathbf{r} \in \mathbb{F}_{q^m}^n} (|\mathcal{B}_\tau(\mathbf{r}) \cap \mathcal{G}|) \leq \sum_{t=\lfloor (d-1)/2 \rfloor + 1}^{\tau} \frac{\begin{bmatrix} n \\ 2t+1-d \end{bmatrix}}{\begin{bmatrix} t \\ 2t+1-d \end{bmatrix}} \tag{3}$$

$$\leq 4 \sum_{t=\lfloor (d-1)/2 \rfloor + 1}^{\tau} q^{(2t-d+1)(n-t)} \leq 4\left(\tau - \left\lfloor \tfrac{d-1}{2} \right\rfloor\right) \cdot q^{(2\tau-d+1)(n-\lfloor (d-1)/2 \rfloor - 1)}. \tag{4}$$

Proof. We consider all words in $\mathbb{F}_{q^m}^n$ with $n \leq m$, therefore these words can be seen as matrices in an n-dimensional space. For any t, where $\lfloor (d-1)/2 \rfloor \leq t \leq d$, there are $\begin{bmatrix} n \\ 2t-d+1 \end{bmatrix}$ subspaces of dimension $2t - d + 1$. Let $\mathbf{r}$ be any fixed word in $\mathbb{F}_{q^m}^n$ and all codewords in $\mathcal{L}$ have $\mathrm{rk}(\mathbf{r} - \mathbf{c}_i) \leq \tau$. Each $\mathbf{r} - \mathbf{c}_i$, for $i = 1, \dots, \ell$, of rank $t \leq \tau$ contains $\begin{bmatrix} t \\ 2t-d+1 \end{bmatrix}$ subspaces of dimension $2t - d + 1$.

Due to Lemma 1, different $\mathbf{r}-\mathbf{c}_i$ have no $(2t-d+1)$-dimensional subspace in common and therefore there are at most $\left[{n \atop 2t-d+1}\right] / \left[{t \atop 2t+1-d}\right]$ possible codewords in rank distance t to the word $\mathbf{r}$. We sum this up for t from $\lfloor (d-1)/2 \rfloor + 1$ up to τ and we obtain (3).

With the bounds for the Gaussian binomial and since $\lfloor (d-1)/2 \rfloor + 1 \leq t \leq \tau$, the upper bound from (4) follows. □

Note that for the special case $\tau = d/2$ and even minimum distance d, the upper bound from (3) is the bound from [3, Equation (4)], which is

$$\ell \leq (q^n-1)\frac{q^{n-d/2}-q^{n-d}}{q^n-2q^{n-d/2}+q^{n-d}} = (q^n-1)\frac{q^{-\tau}-q^{-2\tau}}{1-2q^{-\tau}+q^{-2\tau}} = \frac{q^n-1}{q^\tau-1} = \frac{\left[{n \atop 1}\right]}{\left[{\tau \atop 1}\right]}.$$

Thus, we have proved an upper bound on the maximum list size of a Gabidulin code. Unfortunately, this upper bound is exponential in $n \leq m$ for any $\tau > \lfloor (d-1)/2 \rfloor$ and therefore does not provide any conclusion if polynomial-time list decoding is possible or not in the region up to the Johnson bound.

References

[1] E. M. Gabidulin, "Theory of Codes with Maximum Rank Distance," *Probl. Peredachi Inf.*, vol. 21, no. 1, pp. 3–16, 1985.

[2] C. Faure, "Average Number of Gabidulin Codewords within a Sphere," in *Tenth International Workshop on Algebraic and Combinatorial Coding Theory*, Sept. 2006, pp. 86–89.

[3] D. Augot and P. Loidreau, "Johnson-like bounds for the rank metric," *preprint*, 2011.

[4] O. Ore, "On a Special Class of Polynomials," *Transactions of the American Mathematical Society*, vol. 35, pp. 559–584, 1933.

[5] J. Justesen and T. Hoholdt, "Bounds on list decoding of MDS codes," *Information Theory, IEEE Transactions on*, vol. 47, no. 4, pp. 1604–1609, May 2001.

[6] E. Ben-Sasson, S. Kopparty, and J. Radhakrishnan, "Subspace Polynomials and Limits to List Decoding of Reed–Solomon Codes," *Information Theory, IEEE Transactions on*, vol. 56, no. 1, pp. 113–120, Jan. 2010.

[7] E. R. Berlekamp, *Algebraic Coding Theory, Revised Edition (M-6) (No. M-6)*, revised ed. Aegean Park Pr, June 1984.