On the Decoder Error Probability of Bounded Rank-Distance Decoders for Maximum Rank Distance Codes

Maximilien Gadouleau, *Student Member, IEEE*, and
Zhiyuan Yan, *Member, IEEE*

***Abstract*—In this correspondence, we first introduce the concept of *elementary linear subspace*, which has similar properties to those of a set of coordinates. We then use elementary linear subspaces to derive properties of maximum rank distance (MRD) codes that parallel those of maximum distance separable codes. Using these properties, we show that, for MRD codes with error correction capability t, the decoder error probability of bounded rank distance decoders *decreases exponentially* with t^2 based on the assumption that all errors with the same rank are equally likely.**

***Index Terms*—Bounded distance decoder, decoder error probability, rank metric codes.**

I. Introduction

Although the rank of a matrix has long been known to be a metric [1], the rank metric was first considered for error control codes (ECCs) by Delsarte [2]. ECCs with the rank metric [3]–[6] have been receiving growing attention due to their applications in storage systems [4], public-key cryptosystems [5], space–time coding [6], and network coding [7], [8].

The pioneering works in [2]–[4] have established many important properties of rank metric codes. Independently in [2]–[4], a Singleton bound (up to some variations) on the minimum rank distance of codes was established, and a class of codes that achieve the bound with equality was constructed. We refer to *linear or nonlinear* codes that attain the Singleton bound as maximum rank distance (MRD) codes, and the class of linear MRD codes proposed in [3] as Gabidulin codes henceforth. Different decoding algorithms for Gabidulin codes were proposed in [3], [4], [9], and [10].

In this correspondence, we investigate the error performance of *bounded rank distance decoder* for MRD codes. A bounded rank distance decoder for MRD codes with error correction capability t is guaranteed to correct all errors with rank no more than t. Given a received word, a bounded rank distance decoder either provides an estimate for the transmitted codeword or declares decoder failure. A decoder error occurs when the estimate is not the actual transmitted codeword. The main results of this correspondence are new upper bounds on the decoder error probability (DEP) of bounded rank distance decoders for MRD codes. We emphasize that the DEP considered herein is *conditional*: it is the probability that a bounded rank distance decoder, correcting up to t rank errors, makes an erroneous correction, given that an error with a fixed rank was made. Our bounds indicate that the DEP of MRD codes with error correction capability t *decreases exponentially* with t^2. To derive our bounds, we assume all errors with the same rank are equally likely.

We provide the following remarks on our results.

1) Since decoder failures can be remedied by error masking or retransmission, decoder errors are more detrimental to the overall performance and hence often considered separately (see [11]). This is the main reason we focus on DEP.
2) Note that bounded rank distance decoders guarantee to correct errors with rank up to t. In [12], it was shown that with Gabidulin codes errors with rank beyond t can be corrected when errors occur from the same vector space. However, we do not consider the decoders in [12] and focus on bounded rank distance decoders instead.
3) Our bounds are analogous to the upper bounds on the error probability of bounded Hamming distance decoders for maximum distance separable (MDS) codes in [11] (see [13]–[15] for related results).

We are able to derive our bounds based on an approach which parallels the one in [11]. This was made possible by the concept of *elementary linear subspace* (ELS), which has similar properties to those of a set of coordinates. Using elementary linear subspaces, we also derive useful properties of MRD codes which parallel those of MDS codes. Although our results may be derived without the concept of ELS, we have adopted it in this correspondence since it enables readers to easily relate our approach and results to their counterparts for Hamming metric codes.

The rest of the correspondence is organized as follows. Section II gives a brief review of the rank metric, Singleton bound, and MRD codes. In Section III, we derive some combinatorial properties which are used in the derivation of our upper bounds. In Section IV, we first introduce the concept of elementary linear subspace and study its properties, and then obtain some important properties of MRD codes. In Section V, we derive our upper bounds on the DEP of MRD codes.

II. Preliminaries

Consider an n-dimensional vector $\mathbf{x} = (x_0, x_1, \ldots, x_{n-1}) \in \mathrm{GF}(q^m)^n$. Assume $\{\alpha_0, \alpha_1, \ldots, \alpha_{m-1}\}$ is a basis of $\mathrm{GF}(q^m)$ over $\mathrm{GF}(q)$, then for $j = 0, 1, \ldots, n-1$, x_j can be expanded to an m-dimensional column vector $(x_{0,j}, x_{1,j}, \ldots, x_{m-1,j})^T$ over $\mathrm{GF}(q)$ with respect to the basis $\{\alpha_0, \alpha_1, \ldots, \alpha_{m-1}\}$. Let $\mathbf{X}$ be the $m \times n$ matrix obtained by expanding all the coordinates of $\mathbf{x}$. That is, $\mathbf{X} = \{x_{i,j}\}_{i,j=0}^{m-1,n-1}$ where $x_j = \sum_{i=0}^{m-1} x_{i,j}\alpha_i$. The *rank norm* of the vector $\mathbf{x}$ (over $\mathrm{GF}(q)$), denoted as $\mathrm{rk}(\mathbf{x})$, is defined as $\mathrm{rk}(\mathbf{x}) \overset{\mathrm{def}}{=} \mathrm{rank}(\mathbf{X})$ [3]. The rank norm of $\mathbf{x}$ is also the *maximum* number of coordinates in $\mathbf{x}$ that are linearly independent over $\mathrm{GF}(q)$. The field $\mathrm{GF}(q^m)$ may be viewed as a vector space over $\mathrm{GF}(q)$. The coordinates of $\mathbf{x}$ thus span a linear subspace of $\mathrm{GF}(q^m)$, denoted as $\mathfrak{S}(\mathbf{x})$, such that $\dim(\mathfrak{S}(\mathbf{x})) = \mathrm{rk}(\mathbf{x})$. For all $\mathbf{x}, \mathbf{y} \in \mathrm{GF}(q^m)^n$, it is easily verified that $d(\mathbf{x}, \mathbf{y}) \overset{\mathrm{def}}{=} \mathrm{rk}(\mathbf{x} - \mathbf{y})$ is a metric over $\mathrm{GF}(q^m)^n$, referred to as the *rank metric* henceforth [3]. Hence, the *minimum rank distance* d_{R} of a code is simply the minimum rank distance over all possible pairs of distinct codewords. A code with a minimum rank distance d_{R} can correct all errors with rank up to $t = \lfloor (d_{\mathrm{R}} - 1)/2 \rfloor$.

The minimum rank distance d_{R} of a code of length n over $\mathrm{GF}(q^m)$ satisfies $d_{\mathrm{R}} \leq d_{\mathrm{H}}$ [3], where d_{H} is the minimum Hamming distance of the same code. Due to the Singleton bound on the minimum Hamming distance of block codes [16], the minimum rank distance of a block code of length n and cardinality M over $\mathrm{GF}(q^m)$ thus satisfies $d_{\mathrm{R}} \leq n - \log_{q^m} M + 1$. In this correspondence, we refer to this bound as the Singleton bound for rank metric codes, and to codes that attain the equality as MRD codes. Note that although an MRD code is not

Manuscript received December 8, 2006; revised November 19, 2007. This work was supported in part by Thales Communications Inc. and in part by a grant from the Commonwealth of Pennsylvania, Department of Community and Economic Development, through the Pennsylvania Infrastructure Technology Alliance (PITA). The material in this correspondence was presented in part at IEEE Information Theory Workshop, Chengdu, China, 2006.

The authors are with the Department of Electrical and Computer Engineering, Lehigh University, Bethlehem, PA, 18015 USA (e-mail: magc@lehigh.edu; yan@lehigh.edu).

Communicated by L. M. G. M. Tolhuizen, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2008.924697

necessarily linear, this bound implies that its cardinality is a power of q^m.

The number of vectors of rank $0 \leq u \leq \min\{m, n\}$ in GF$(q^m)^n$ is given by $N_u = \begin{bmatrix} n \\ u \end{bmatrix} A(m,u)$, where $A(m,u)$ is defined as follows: $A(m,0) = 1$ and $A(m,u) = \prod_{i=0}^{u-1}(q^m - q^i)$ for $u \geq 1$. The $\begin{bmatrix} n \\ u \end{bmatrix}$ term is the Gaussian binomial [17], defined as $\begin{bmatrix} n \\ u \end{bmatrix} = A(n,u)/A(u,u)$. Note that $\begin{bmatrix} n \\ u \end{bmatrix}$ is the number of u-dimensional linear subspaces of GF$(q)^n$ [17].

Note that following the approach in [3], the vector form over GF(q^m) is used to represent rank metric codes although their rank weight is defined by their corresponding $m \times n$ code matrices over GF(q). Naturally, rank metric codes can be studied in the matrix form (see [2], [4]). The vector form is chosen in this correspondence since our results and their derivations for rank metric codes can be related to their counterparts for Hamming metric codes.

III. Combinatorial Results

In this section, we derive some combinatorial properties which will be instrumental in the derivation of our results in Section V.

Lemma 1: For $0 \leq u \leq m$, $K_q q^{mu} < A(m,u) \leq q^{mu}$, where $K_q = \prod_{j=1}^{\infty}(1 - q^{-j})$. Also, for $0 \leq u \leq m-1$, $A(m,u) > \frac{q}{q-1} K_q q^{mu}$. Finally, for $0 \leq u \leq \lfloor m/2 \rfloor$, $A(m,u) \geq \frac{q^2-1}{q^2} q^{mu}$.

Proof: The upper bound is trivial. We now prove the lower bounds. We have $A(m,u) = q^{mu}\prod_{i=0}^{u-1}(1 - q^{i-m}) > q^{mu}K_q$. The second lower bound follows from $A(m, m-1) = \frac{q}{q-1} q^{-m} A(m,m)$.

The third lower bound clearly holds for $m = 0$ and $m = 1$. Let us assume $m \geq 2$ henceforth and denote $\frac{q^{mu}}{A(m,u)}$ as $D(m,u)$. It can be easily verified that $D(m,u)$ is an increasing function of u. Thus, it suffices to show that $D(m, \lfloor m/2 \rfloor) \leq \frac{q^2}{q^2-1}$ for $m \geq 2$. First, if m is odd, $m = 2p+1$, it can be easily shown that $D(2p+1, p) < D(2p, p)$. Hence, we need to consider only the case where $m = 2p$, with $p \geq 1$. Let us further show that $D(2p, p)$ is a monotonically decreasing function of p since

$$\begin{aligned} &D(2p+2, p+1) \\ &= \frac{q^{2p+2}}{q^{2p+2}-1} \cdot \frac{q^{2p+2}}{q^{2p+2}-q} \cdot \frac{q^{2p+2}-q^{p+1}}{q^{2p+2}} D(2p,p) \\ &= \frac{q^{2p+1}(q^{2p+2}-q^{p+1})}{(q^{2p+1}-1)(q^{2p+2}-1)} D(2p,p). \end{aligned}$$

The maximum of $D(2p,p)$ is hence given by $D(2,1) = \frac{q^2}{q^2-1}$. □

It is worth noting that K_q above represents the fraction of invertible $m \times m$ matrices over GF(q) as m approaches infinity, and that K_q increases with q.

Corollary 1: For $0 \leq t \leq n$, we have $\begin{bmatrix} n \\ t \end{bmatrix} < K_q^{-1} q^{t(n-t)}$.

Proof: By definition, $\begin{bmatrix} n \\ t \end{bmatrix} = A(n,t)/A(t,t)$. Since $A(n,t) \leq q^{nt}$ and by Lemma 1, $A(t,t) > K_q q^{t^2}$, we obtain $\begin{bmatrix} n \\ t \end{bmatrix} < K_q^{-1} q^{t(n-t)}$. □

IV. Properties of MRD Codes

Many properties of MDS codes are established by studying sets of coordinates. These sets of coordinates may be viewed as linear subspaces which have a basis of vectors with Hamming weight 1. Similarly, some properties of MRD codes may be established using elementary linear subspaces (ELS's), which can be considered as the counterparts of sets of coordinates.

A. Elementary Linear Subspaces

It is a well-known fact in linear algebra (see, for example, [3]) that a vector $\mathbf{x}$ of rank $\text{rk}(\mathbf{x}) \leq u$ can be represented as $\mathbf{x} = (x_0, x_1, \ldots, x_{n-1}) = (e_0, e_1, \ldots, e_{u-1})\mathbf{A}$, where $e_j \in$ GF(q^m) for $j = 0, 1, \ldots, u-1$ and $\mathbf{A}$ is a $u \times n$ matrix over GF(q) of full rank u. The concept of elementary linear subspace can be introduced as a consequence of this representation. However, due to its usefulness in our approach we define the concept formally and study its properties below from a different perspective.

Definition 1 (Elementary Linear Subspace): A linear subspace $\mathcal{V}$ of GF$(q^m)^n$ is said to be elementary if it has a basis B consisting of row vectors in GF$(q)^n$. B is called an elementary basis of $\mathcal{V}$. For $0 \leq v \leq n$, we define $E_v(q^m, n)$ as the set of all ELS's with dimension v in GF$(q^m)^n$.

By definition, a linear subspace $\mathcal{V}$ with dimension v is an ELS if and only if it is the row span of a $v \times n$ matrix $\mathbf{B}$ over GF(q) with full rank. Thus there exists a bijection between $E_v(q^m, n)$ and $E_v(q, n)$, and $|E_v(q^m, n)| = \begin{bmatrix} n \\ v \end{bmatrix}$. Also, it can be easily shown that a linear subspace $\mathcal{V}$ of GF$(q^m)^n$ is an ELS if and only if there exists a basis consisting of vectors of rank 1for $\mathcal{V}$.

Next, we show that the properties of ELS's are similar to those of sets of coordinates.

Proposition 1: For all $\mathcal{V} \in E_v(q^m, n)$ there exists $\bar{\mathcal{V}} \in E_{n-v}(q^m, n)$ such that $\mathcal{V} \oplus \bar{\mathcal{V}} =$ GF$(q^m)^n$, where $\mathcal{V} \oplus \bar{\mathcal{V}}$ denotes the direct sum of $\mathcal{V}$ and $\bar{\mathcal{V}}$.

Proof: Clearly, the ELS $\bar{\mathcal{V}}$ having elementary basis $\bar{B}$ such that $B \cup \bar{B}$ is a basis of GF$(q)^n$ satisfies $\mathcal{V} \oplus \bar{\mathcal{V}} =$ GF$(q^m)^n$. □

We say that $\bar{\mathcal{V}}$ is an *elementary complement* of $\mathcal{V}$. Even though an elementary complement always exists, we remark that it may not be unique.

The diameter of a code for the Hamming metric is defined in [16] as the maximum Hamming distance between two codewords. Similarly, we can define the rank diameter of a linear subspace.

Definition 2: The rank diameter of a linear subspace $\mathcal{L}$ of GF$(q^m)^n$ is defined to be the maximum rank among the vectors in $\mathcal{L}$, i.e., $\delta(\mathcal{L}) \overset{\text{def}}{=} \max_{\mathbf{x} \in \mathcal{L}}\{\text{rk}(\mathbf{x})\}$.

Proposition 2: For all $\mathcal{V} \in E_v(q^m, n)$, $\delta(\mathcal{V}) \leq v$. Furthermore, if $v \leq m$, then $\delta(\mathcal{V}) = v$.

Proof: Any vector $\mathbf{x} \in \mathcal{V}$ can be expressed as the sum of at most v vectors of rank 1, hence its rank is upper bounded by v. Thus, $\delta(\mathcal{V}) \leq v$ by Definition 2. If $v \leq m$, we show that there exists a vector in $\mathcal{V}$ with rank v. Let $B = \{\mathbf{b}_i\}_{i=0}^{v-1}$ be an elementary basis of $\mathcal{V}$, and consider $\mathbf{y} = \sum_{i=0}^{v-1} \alpha_i \mathbf{b}_i$, where $\{\alpha_i\}_{i=0}^{m-1}$ is a basis of GF(q^m) over GF(q). If we expand the coordinates of $\mathbf{y}$ with respect to the basis $\{\alpha_i\}_{i=0}^{m-1}$, we obtain $\mathbf{Y} = \left(\mathbf{b}_0^T, \ldots, \mathbf{b}_{v-1}^T, \mathbf{0}^T, \ldots, \mathbf{0}^T\right)^T$. Since the row vectors $\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_{v-1}$ are linearly independent over GF(q), $\mathbf{Y}$ has rank v and $\text{rk}(\mathbf{y}) = v$. □

Lemma 2: A vector $\mathbf{x} \in$ GF$(q^m)^n$ has rank $\leq u$ if and only if it belongs to some $\mathcal{A} \in E_u(q^m, n)$.

Proof: The necessity is obvious. We now prove the sufficiency. Suppose $\mathbf{x} = (x_0, x_1, \ldots, x_{n-1})$ has rank u, and without loss of generality, assume its first u coordinates are linearly independent. Thus, for $j = 0, \ldots, n-1$, $x_j = \sum_{i=0}^{u-1} a_{ij} x_i$, where $a_{ij} \in$ GF(q). That is, $\mathbf{x} = (x_0, x_1, \ldots, x_{u-1})\mathbf{A}$, where $\mathbf{A} = \{a_{ij}\}_{i=0, j=0}^{u-1, n-1} = (\mathbf{a}_0^T, \ldots, \mathbf{a}_{u-1}^T)^T$. Thus $\mathbf{x} = \sum_{i=0}^{u-1} x_i \mathbf{a}_i$, with $\mathbf{a}_i \in$ GF$(q)^n$ for $0 \leq i \leq u-1$. Let $\mathcal{A}$ be the ELS of GF$(q^m)^n$ spanned by $\mathbf{a}_i$'s, then $\dim(\mathcal{A}) = u$ and $\mathbf{x} \in \mathcal{A}$. This proof can be easily adapted to the case where $\text{rk}(\mathbf{x}) < u$. □

Let $\mathcal{L}$ be a linear subspace of GF$(q^m)^n$ and let $\bar{\mathcal{L}}$ be complementary to $\mathcal{L}$, i.e., $\mathcal{L} \oplus \bar{\mathcal{L}} = \text{GF}(q^m)^n$. We denote the projection of $\mathbf{x}$ on $\mathcal{L}$ along $\bar{\mathcal{L}}$ as $\mathbf{x}_{\mathcal{L}}$ [18]. Remark that $\mathbf{x} = \mathbf{x}_{\mathcal{L}} + \mathbf{x}_{\bar{\mathcal{L}}}$. Note that for any given linear subspace $\mathcal{L}$, its complementary linear subspace $\bar{\mathcal{L}}$ is not unique. Thus, $\mathbf{x}_{\mathcal{L}}$ depends on both $\mathcal{L}$ and $\bar{\mathcal{L}}$, and is well-defined only when both $\mathcal{L}$ and $\bar{\mathcal{L}}$ are given. All the projections in this correspondence are with respect to a pair of fixed linear subspaces complementary to each other.

Definition 3: Let $\mathbf{x} \in \text{GF}(q^m)^n$ and $\mathcal{L}$ be a linear subspace. The vector $\mathbf{x}$ *vanishes* on $\mathcal{L}$ if there exists a linear subspace $\bar{\mathcal{L}}$ complementary to $\mathcal{L}$ such that $\mathbf{x} = \mathbf{x}_{\bar{\mathcal{L}}}$.

Lemma 3: A vector $\mathbf{x} \in \text{GF}(q^m)^n$ has rank $\leq u$ if and only if it vanishes on some $\mathcal{B} \in E_{n-u}(q^m, n)$.

Proof: Suppose $\mathbf{x}$ has rank $\leq u$. By Lemma 2, there exists $\mathcal{A} \in E_u(q^m, n)$ such that $\mathbf{x} \in \mathcal{A}$. Let $\bar{\mathcal{A}}$ be an elementary complement of $\mathcal{A}$. Thus, $\mathbf{x}$ vanishes on $\bar{\mathcal{A}}$ by definition. Also, suppose $\mathbf{x}$ vanishes on an ELS $\bar{\mathcal{B}}$ with dimension greater than $n-u$. Then there exists an ELS $\mathcal{B}$ with dimension $< u$ such that $\mathbf{x} \in \mathcal{B}$, which contradicts Lemma 2. □

The decomposition over two complementary ELS's induces a mapping from GF$(q^m)^n$ to GF$(q^m)^v$.

Definition 4: Let $\mathcal{V} \in E_v(q^m, n)$ be the row span of $\mathbf{B}$ with an elementary complement $\bar{\mathcal{V}}$. For any $\mathbf{x} \in \text{GF}(q^m)^n$, we define $r_{\mathcal{V}}(\mathbf{x}) = (r_0, \ldots, r_{v-1}) \in \text{GF}(q^m)^v$ to be $r_{\mathcal{V}}(\mathbf{x}) = \mathbf{x}_{\mathcal{V}}\mathbf{B}^{-R}$, where $\mathbf{B}^{-R}$ is the right inverse of $\mathbf{B}$.

We remark that the $r_{\mathcal{V}}$ function is linear and since $\mathbf{B}^{-R}$ has full rank, we have $\text{rk}(r_{\mathcal{V}}(\mathbf{x})) = \text{rk}(\mathbf{x}_{\mathcal{V}})$ for all $\mathbf{x}$.

Definition 5: For $\mathcal{V} \in E_v(q^m, n)$ the row span of $\mathbf{B}$, let $\bar{\mathcal{V}}$ be an elementary complement of $\mathcal{V}$ which is the row span of $\bar{\mathbf{B}}$. For any $\mathbf{x} \in \text{GF}(q^m)^n$, we define $s_{\mathcal{V},\bar{\mathcal{V}}}(\mathbf{x}) = (r_{\mathcal{V}}(\mathbf{x}), r_{\bar{\mathcal{V}}}(\mathbf{x})) \in \text{GF}(q^m)^n$.

Lemma 4: For all $\mathbf{x} \in \text{GF}(q^m)^n$, $\text{rk}(s_{\mathcal{V},\bar{\mathcal{V}}}(\mathbf{x})) = \text{rk}(\mathbf{x})$.

Proof: Note that $\mathbf{x} = s_{\mathcal{V},\bar{\mathcal{V}}}(\mathbf{x})\hat{\mathbf{B}}$ where $\hat{\mathbf{B}} = (\mathbf{B}, \bar{\mathbf{B}})^T$ is an $n \times n$ matrix over GF(q) with full rank. Therefore $\text{rk}(s_{\mathcal{V},\bar{\mathcal{V}}}(\mathbf{x})) = \text{rk}(\mathbf{x})$. □

Corollary 2: For all $\mathbf{x} \in \text{GF}(q^m)^n$ and two complementary ELS's $\mathcal{V}$ and $\bar{\mathcal{V}}$, $0 \leq \text{rk}(\mathbf{x}_{\mathcal{V}}) \leq \text{rk}(\mathbf{x})$ and $\text{rk}(\mathbf{x}) \leq \text{rk}(\mathbf{x}_{\mathcal{V}}) + \text{rk}(\mathbf{x}_{\bar{\mathcal{V}}})$.

It can be easily shown that the second inequality in Corollary 2 can be strict in some cases. For example, consider $\mathbf{x} = (1, 1) \in \text{GF}(q^2)^2$. For appropriate ELS's $\mathcal{V}$ and $\bar{\mathcal{V}}$, $\mathbf{x}_{\mathcal{V}} = (1, 0)$ and $\mathbf{x}_{\bar{\mathcal{V}}} = (0, 1)$. Clearly, $\text{rk}(\mathbf{x}) < \text{rk}(\mathbf{x}_{\mathcal{V}}) + \text{rk}(\mathbf{x}_{\bar{\mathcal{V}}})$. However, when $\mathcal{V}$ and $\bar{\mathcal{V}}$ are two complementary sets of coordinates, the Hamming weight of any vector $\mathbf{x} \in \text{GF}(q^m)^n$ is the sum of the Hamming weights of the projections of $\mathbf{x}$ on $\mathcal{V}$ and $\bar{\mathcal{V}}$. Therefore, Corollary 2 illustrates the difference between ELS's and sets of coordinates.

B. Properties of MRD Codes

We now derive some useful properties of MRD codes, which will be instrumental in Section V. These properties are similar to those of MDS codes. Let C be an MRD code over GF(q^m) with length n $(n \leq m)$, cardinality q^{mk}, redundancy $r = n - k$, and minimum rank distance $d_{\text{R}} = n - k + 1$. We emphasize that C may be *linear or nonlinear*, which is necessary for our derivation in Section V. First, we derive the basic combinatorial property of MRD codes.

Lemma 5 (Basic Combinatorial Property): For any $\mathcal{K} \in E_k(q^m, n)$ and its elementary complement $\bar{\mathcal{K}}$ and any vector $\mathbf{k} \in \mathcal{K}$, there exists a unique codeword $\mathbf{c} \in C$ such that $\mathbf{c}_{\mathcal{K}} = \mathbf{k}$.

Proof: Suppose there exist $\mathbf{c}, \mathbf{d} \in C, \mathbf{c} \neq \mathbf{d}$ such that $\mathbf{c}_{\mathcal{K}} = \mathbf{d}_{\mathcal{K}}$. Then $\mathbf{c} - \mathbf{d} \in \bar{\mathcal{K}}$, and $0 < \text{rk}(\mathbf{c} - \mathbf{d}) \leq n - k$ by Proposition 2, which contradicts the fact that C is MRD. Then all the codewords lead to different projections on $\mathcal{K}$. Since $|C| = |\mathcal{K}| = q^{mk}$, for any $\mathbf{k} \in \mathcal{K}$ there exists a unique $\mathbf{c}$ such that $\mathbf{c}_{\mathcal{K}} = \mathbf{k}$. □

Lemma 5 allows us to bound the rank distribution of MRD codes.

Lemma 6 (Bound on the Rank Distribution): Let A_u be the number of codewords in C with rank u. Then, for $u \geq d_{\text{R}}$

$$A_u \leq \begin{bmatrix} n \\ u \end{bmatrix} A(m, u - r). \tag{1}$$

Proof: By Lemma 3, any codeword $\mathbf{c}$ with rank $u \geq d_{\text{R}}$ vanishes on an ELS with dimension $v = n - u$. Thus (1) can be established by first determining the number of codewords vanishing on a given ELS of dimension v, and then multiplying by the number of such ELS's, $\begin{bmatrix} n \\ v \end{bmatrix} = \begin{bmatrix} n \\ u \end{bmatrix}$. For $\mathcal{V} \in E_v(q^m, n)$, $\mathcal{V}$ is properly contained in an ELS $\mathcal{K}$ with dimension k since $v \leq k - 1$. By Lemma 5, $\mathbf{c}$ is completely determined by $\mathbf{c}_{\mathcal{K}}$. Given an elementary basis of $\mathcal{K}$ having v elements that span $\mathcal{V}$, it suffices to determine $r_{\mathcal{K}}(\mathbf{c})$. However, $\mathbf{c}_{\mathcal{K}}$ vanishes on $\mathcal{V}$, hence v of the coordinates of $r_{\mathcal{K}}(\mathbf{c})$ must be zero. By Lemma 3, the other $k - v$ coordinates must be nonzero, and since $\text{rk}(\mathbf{c}_{\mathcal{K}}) = n - v$, these coordinates must be linearly independent. Hence, a codeword that vanishes on $\mathcal{V}$ is completely determined by $k - v$ arbitrary linearly independent coordinates. There are at most $A(m, k - v) = A(m, u - r)$ choices for these coordinates, and hence at most $A(m, u - r)$ codewords that vanish on $\mathcal{V}$. □

Note that the exact formula for the rank distribution of *linear* MRD codes was derived independently in [2] and [3]. Thus, tighter bounds on A_u can be derived for linear codes. However, our derivation of the DEP of MRD codes in Section V requires bounds on A_u for both linear and nonlinear MRD codes. Therefore, the exact rank distribution of linear MRD codes cannot be used, and the bound in (1) should be used instead.

Definition 6 (Restriction of a Code): For $\mathcal{V} \in E_v(q^m, n)$ $(k \leq v \leq n)$ with elementary basis B and its elementary complement $\bar{\mathcal{V}}$, $C_{\mathcal{V}} = \{r_{\mathcal{V}}(\mathbf{c}) | \mathbf{c} \in C\}$ is called the restriction of C to $\mathcal{V}$.

It is well known that a punctured MDS code is an MDS code [16]. We now show that the restriction of an MRD code to an ELS is also MRD.

Lemma 7 (Restriction of an MRD Code): For all ELS $\mathcal{V}$ with dimension v $(k \leq v \leq n)$, $C_{\mathcal{V}}$ is an MRD code with length v, cardinality q^{mk}, and minimum rank distance $d_{\text{R}} = v - k + 1$ over GF(q^m).

Proof: For $\mathbf{c} \neq \mathbf{d} \in C$, consider $\mathbf{x} = \mathbf{c} - \mathbf{d}$. By the property of $r_{\mathcal{V}}$ function, we have

$$\begin{aligned} \text{rk}(r_{\mathcal{V}}(\mathbf{c}) - r_{\mathcal{V}}(\mathbf{d})) &= \text{rk}(r_{\mathcal{V}}(\mathbf{c} - \mathbf{d})) = \text{rk}(\mathbf{x}_{\mathcal{V}}) \\ &\geq \text{rk}(\mathbf{x}) - \text{rk}(\mathbf{x}_{\bar{\mathcal{V}}}) \geq n - k + 1 - (n - v) = v - k + 1. \end{aligned}$$

Therefore, $C_{\mathcal{V}}$ is a code over GF(q^m) with length v, cardinality q^{mk}, and minimum rank distance $\geq v - k + 1$. The Singleton bound on $C_{\mathcal{V}}$ completes the proof. □

V. Performance of MRD Codes

We evaluate the error performance of MRD codes using a bounded rank distance decoder. We assume that the errors are additive and that all errors with the same rank are equiprobable. A bounded rank distance decoder produces a codeword within rank distance $t = \lfloor (d_{\text{R}} - 1)/2 \rfloor$ of the received word if it can find one, and declares a decoder failure if it cannot. In the following, we first derive bounds on the DEP assuming the error has rank u. In the end, we derive a bound on the DEP that does *not* depend on u. We denote the probabilities of decoder error and failure for the bounded rank distance decoder — for error correction capability t and an error of rank u — as $P_E(t;u)$ and $P_F(t;u)$, respectively. Clearly, $P_F(t;u) = P_E(t;u) = 0$ for $u \leq t$ and $P_E(t;u) = 0$ and $P_F(t;u) = 1$ for $t < u < d_{\mathrm{R}} - t$, which occurs only if $d_{\mathrm{R}} = 2t+2$. Thus we investigate the case where $u \geq d_{\mathrm{R}} - t$ and $P_E(t;u)$ characterizes the performance of the code, as $P_E(t;u) + P_F(t;u) = 1$.

Since our derivation below is transparent to the transmitted codeword, we assume without loss of generality that the all-zero vector is a codeword and is transmitted. Thus, the received word can be any vector with rank u with equal probability. We call a vector decodable if it lies within rank distance t of some codeword. If D_u denotes the number of decodable vectors of rank u, then for $u \geq d_{\mathrm{R}} - t$ we have

$$P_E(t;u) = \frac{D_u}{N_u} = \frac{D_u}{\begin{bmatrix}n\\u\end{bmatrix} A(m,u)}. \tag{2}$$

Hence, the main challenge is to derive upper bounds on D_u. We consider two cases, $u \geq d_{\mathrm{R}}$ and $d_{\mathrm{R}} - t \leq u < d_{\mathrm{R}}$, separately.

Proposition 3: For $u \geq d_{\mathrm{R}}$, then $D_u \leq \begin{bmatrix}n\\u\end{bmatrix} A(m, u-r) V_t$, where $V_t = \sum_{i=0}^{t} N_i$ is the volume of a ball of rank radius t.

Proof: Any decodable vector can be uniquely written as $\mathbf{c} + \mathbf{e}$, where $\mathbf{c} \in C$ and $\mathrm{rk}(\mathbf{e}) \leq t$. For a fixed $\mathbf{e}$, $C + \mathbf{e}$ is an MRD code, which satisfies (1). Therefore, the number of decodable words of rank u is at most $\begin{bmatrix}n\\u\end{bmatrix} A(m, u-r)$, by Lemma 6, multiplied by the number of error vectors, V_t. □

Lemma 8: Suppose $\mathbf{y} = (y_0, \ldots, y_{v-1}) \in \mathrm{GF}(q^m)^v$ has rank w. Then there exist $\begin{bmatrix}u\\s-w\end{bmatrix} A(m-w, s-w) q^{wu}$ vectors $\mathbf{z} = (z_0, \ldots, z_{u-1}) \in \mathrm{GF}(q^m)^u$ such that $\mathbf{x} = (\mathbf{y}, \mathbf{z}) \in \mathrm{GF}(q^m)^{u+v}$ has rank s.

Proof: Let $\mathfrak{T}$ be an $(m-w)$-dimensional subspace of $\mathrm{GF}(q^m)$ such that $\mathfrak{T} \oplus \mathfrak{S}(\mathbf{y}) = \mathrm{GF}(q^m)$. We can thus express $\mathbf{z}$ as $\mathbf{z} = \mathbf{a} + \mathbf{b}$, where $a_i \in \mathfrak{T}$ and $b_i \in \mathfrak{S}(\mathbf{y})$ for all i. Since $\mathrm{rk}(\mathbf{x}) = \mathrm{rk}(\mathbf{y}) + \mathrm{rk}(\mathbf{a})$, we have $\mathrm{rk}(\mathbf{a}) = s - w$, and hence there are $\begin{bmatrix}u\\s-w\end{bmatrix} A(m-w, s-w)$ possible choices for $\mathbf{a}$. Also, there are q^{wu} choices for the vector $\mathbf{b}$. □

We also obtain a bound similar to the one in Proposition 3 for $d_{\mathrm{R}} - t \leq u < d_{\mathrm{R}}$.

Proposition 4: For $d_{\mathrm{R}} - t \leq u < d_{\mathrm{R}}$, then $D_u < \frac{q^2}{q^2-1} \begin{bmatrix}n\\u\end{bmatrix} (q^m - 1)^{u-r} V_t$.

Proof: Recall that a decodable vector of rank u can be expressed as $\mathbf{c} + \mathbf{e}$, where $\mathbf{c} \in C$ and $\mathrm{rk}(\mathbf{e}) \leq t$. This decodable vector vanishes on an ELS $\mathcal{V}$ with dimension $v = n - u$ by Lemma 3. We have $w \overset{\mathrm{def}}{=} \mathrm{rk}(r_{\mathcal{V}}(\mathbf{c})) \leq t$ by Corollary 2. $C_{\mathcal{V}}$ is an MRD code by Lemma 7, hence $w \geq d_{\mathrm{R}} - u$. A codeword $\mathbf{c} \in C$ is completely determined by $\mathbf{c}_{\mathcal{V}}$ by Lemma 7. Denoting $r' = r - u$, the number of codewords in $C_{\mathcal{V}}$ with rank w is at most $\begin{bmatrix}v\\w\end{bmatrix} A(m, w - r')$ by Lemma 6. For each codeword $\mathbf{c}$ such that $\mathrm{rk}(r_{\mathcal{V}}(\mathbf{c})) = w$, we count the number of error vectors $\mathbf{e}$ such that $r_{\mathcal{V}}(\mathbf{c}) + r_{\mathcal{V}}(\mathbf{e}) = \mathbf{0}$. Suppose that $\mathbf{e}$ has rank s $(w \leq s \leq t)$, then $s_{\mathcal{V},\bar{\mathcal{V}}}(\mathbf{e}) = (-r_{\mathcal{V}}(\mathbf{c}), r_{\bar{\mathcal{V}}}(\mathbf{e}))$ has rank s by Lemma 4. By Lemma 8, there are at most $\begin{bmatrix}u\\s-w\end{bmatrix} A(m-w, s-w) q^{wu}$ choices for $r_{\bar{\mathcal{V}}}(\mathbf{e})$, and hence as many choices for $\mathbf{e}$.

The total number $D_{\mathcal{V}}$ of decodable vectors vanishing on $\mathcal{V}$ is then at most

$$D_{\mathcal{V}} \leq \sum_{w=d_{\mathrm{R}}-u}^{t} \begin{bmatrix}v\\w\end{bmatrix} A(m, w-r') \cdot \sum_{s=w}^{t} \begin{bmatrix}u\\s-w\end{bmatrix} A(m-w, s-w) q^{wu}. \tag{3}$$

We have $A(m, w-r+u) \leq (q^m-1)^{w-r+u}$ and $q^{wu} A(m-w, s-w) \leq q^{w(u-s+w)} A(m, s-w)$. Equation (3) implies

$$\begin{aligned} D_{\mathcal{V}} &\leq (q^m-1)^{u-r} \sum_{s=w}^{t} \sum_{w=d_{\mathrm{R}}-u}^{s} \begin{bmatrix}v\\w\end{bmatrix} \begin{bmatrix}u\\s-w\end{bmatrix} \\ &\quad \cdot q^{w(u-s+w)} A(m, s-w) (q^m-1)^w \\ &< (q^m-1)^{u-r} \sum_{s=d_{\mathrm{R}}-u}^{t} q^{ms} \\ &\quad \cdot \sum_{w=d_{\mathrm{R}}-u}^{s} \begin{bmatrix}v\\w\end{bmatrix} \begin{bmatrix}u\\s-w\end{bmatrix} q^{w(u-s+w)}. \end{aligned}$$

Using [17, p. 225]: $\sum_{w=0}^{s} \begin{bmatrix}v\\w\end{bmatrix} \begin{bmatrix}u\\s-w\end{bmatrix} q^{w(u-s+w)} = \begin{bmatrix}v+u\\s\end{bmatrix}$, we obtain $D_{\mathcal{V}} < (q^m-1)^{u-r} \sum_{s=d_{\mathrm{R}}-u}^{t} q^{ms} \begin{bmatrix}n\\s\end{bmatrix}$. By Lemma 1, we find that

$$\begin{aligned} D_{\mathcal{V}} &< (q^m-1)^{u-r} \sum_{s=d_{\mathrm{R}}-u}^{t} \frac{q^2}{q^2-1} A(m,s) \begin{bmatrix}n\\s\end{bmatrix} \\ &< \frac{q^2}{q^2-1} (q^m-1)^{u-r} V_t. \end{aligned}$$

The result follows by multiplying the bound on $D_{\mathcal{V}}$ by $\begin{bmatrix}n\\v\end{bmatrix}$, the number of ELS's of dimension v. □

Finally, we can derive our bounds on the DEP.

Proposition 5: For $d_{\mathrm{R}} - t \leq u < d_{\mathrm{R}}$, the DEP satisfies

$$P_E(t;u) < \frac{q^2}{q^2-1} \frac{(q^m-1)^{u-r}}{A(m,u)} V_t. \tag{4}$$

For $u \geq d_{\mathrm{R}}$, the DEP satisfies

$$P_E(t;u) < \frac{A(m, u-r)}{A(m,u)} V_t. \tag{5}$$

Proof: The bound in (4) follows directly from (2) and Proposition 4, while the bound in (5) follows directly from (2) and Proposition 3. □

The result may be weakened in order to find a bound on the DEP in exponential form which depends on t only. In order to obtain this bound, we need a bound on V_t first.

Lemma 9: For $0 \leq t \leq \min\{n,m\}$, $V_t \leq \begin{bmatrix}n\\t\end{bmatrix} q^{mt} < K_q^{-1} q^{t(m+n-t)}$.

Proof: Without loss of generality, assume the ball is centered at zero. From Lemma 2, every vector $\mathbf{x}$ in the ball belongs to some ELS $\mathcal{V}$ with dimension t. Since $|\mathcal{V}| = q^{mt}$ and $|E_t(q^m, n)| = \begin{bmatrix}n\\t\end{bmatrix}$, it follows that $V_t \leq \begin{bmatrix}n\\t\end{bmatrix} q^{mt}$. By Corollary 1, we have $\begin{bmatrix}n\\t\end{bmatrix} < K_q^{-1} q^{t(n-t)}$, and hence $V_t < K_q^{-1} q^{t(m+n-t)}$. □

Proposition 6: For $u \geq d_{\mathrm{R}} - t$, the DEP satisfies

$$P_E(t;u) < \frac{q^{-t^2}}{K_q^2}. \tag{6}$$

Proof: First suppose that $u \geq d_{\mathrm{R}}$. Applying $A(m,u) > K_q q^{mu}$ in Lemma 1 and Lemma 9 to (5), we obtain $P_E(t;u) < K_q^{-2} q^{-mr+t(m+n-t)}$. Since $n \leq m$ and $2t \leq r$, it follows that $P_E(t;u) < K_q^{-2} q^{-t^2}$. For $d_{\mathrm{R}} - t \leq u < d_{\mathrm{R}}$, applying $A(m,u) > \frac{q}{q-1} K_q q^{mu}$ in Lemma 1 and Lemma 9 to (4), we obtain $P_E(t;u) < \frac{q^2(q-1)}{q(q^2-1)} K_q^{-2} q^{-mr+t(m+n-t)} < K_q^{-2} q^{-t^2}$. □

Based on the proof above, it is clear that the bound in Proposition 6 is less tight than those in Proposition 5. However, the bound in Proposition 6 does not depend on the rank of the error at all. This implies that the bound applies to any error vector provided the errors with the same rank are equiprobable. Based on conditional probability, we can easily establish the following corollary.

Corollary 3: For an MRD code with $d_{\mathrm{R}} = 2t + 1$ and any additive error such that the errors with the same rank are equiprobable, the DEP of a bounded rank distance decoder satisfies $P_E(t) < K_q^{-2} q^{-t^2}$.

Acknowledgment

The authors are grateful to the anonymous reviewers and the Associate Editor Dr. Ludo Tolhuizen for their constructive comments, which improved both the results and the presentation of the correspondence. They wish to thank one of the reviewers for suggesting the present proofs for Proposition 1 and Lemma 6, which are more concise than their original proofs. Due to another reviewer's comments, the precise expression in Lemma 8 replaced an upper bound in our original submission.

References

[1] L. Hua, "A theorem on matrices over a field and its applications," *Chinese Math. Soc.*, vol. 1, no. 2, pp. 109–163, 1951.

[2] P. Delsarte, "Bilinear forms over a finite field, with applications to coding theory," *J. Combin. Theory A*, vol. 25, pp. 226–241, 1978.

[3] E. M. Gabidulin, "Theory of codes with maximum rank distance," *Probl. Inf. Transm.*, vol. 21, no. 1, pp. 1–12, Jan. 1985.

[4] R. M. Roth, "Maximum-rank array codes and their application to criss-cross error correction," *IEEE Trans. Inf. Theory*, vol. 37, no. 2, pp. 328–336, Mar. 1991.

[5] E. M. Gabidulin, A. V. Paramonov, and O. V. Tretjakov, "Ideals over a non-commutative ring and their application in cryptology," *LNCS*, vol. 573, pp. 482–489, 1991.

[6] P. Lusina, E. M. Gabidulin, and M. Bossert, "Maximum rank distance codes as space-time codes," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2757–2760, Oct. 2003.

[7] R. Koetter and F. R. Kschischang, "Coding for errors and erasures in random network coding," *IEEE Trans. Inf. Theory*, submitted for publication.

[8] D. Silva, F. R. Kschischang, and R. Koetter, "A rank-metric approach to error control in random network coding," *IEEE Trans. Inf. Theory*, submitted for publication.

[9] G. Richter and S. Plass, "Fast decoding of rank-codes with rank errors and column erasures," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2004, p. 398.

[10] P. Loidreau, "A Welch-Berlekamp like algorithm for decoding Gabidulin codes," in *Proc. 4th Int. Workshop Coding Crypto.*, 2005.

[11] R. J. McEliece and L. Swanson, "On the decoder error probability for Reed-Solomon codes," *IEEE Trans. Inf. Theory*, vol. 32, no. 5, pp. 701–703, Sep. 1986.

[12] P. Loidreau and R. Overbeck, "Decoding rank errors beyond the error-correction capability," *Proc. ACCT-10*, 2006.

[13] L. Driessen and L. Vries, "Performance calculation of the compact disc error correcting code," in *Proc. Int. Conf. Video Data Recording*, Apr. 1982, pp. 385–395.

[14] K.-M. Cheung, "More on the decoder error probability for Reed-Solomon codes," *IEEE Trans. Inf. Theory*, vol. 35, no. 4, pp. 895–900, Jul. 1989.

[15] L. Tolhuizen, "A universal upper bound on the miscorrection probability with bounded distance decoding for a code used on an error-value symmetric channel," in *Proc. Eurocode, Int. Symp. Coding Theory and Appl.*, 1992, pp. 313–320.

[16] F. MacWilliams and N. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[17] G. E. Andrews, *The Theory of Partitions*, G.-C. Rota, Ed. Reading, MA: Addison-Wesley, 1976, vol. 2, Encyclopedia of Mathematics and its Applications.

[18] S. Roman, *Advanced Linear Algebra*, 2nd ed. New York: Springer, 2005, vol. 135, Graduate Texts in Mathematics.

On the Probability Distribution of Superimposed Random Codes

Bernd Günther

***Abstract*—In this correspondence, a systematic study of the probability distribution of superimposed random codes is presented through the use of generating functions. Special attention is paid to the cases of either uniformly distributed but not necessarily independent or nonuniform but independent bit structures. Recommendations for optimal coding strategies are derived.**

***Index Terms*—Database indexing, false drop estimates, generating functions, probability distribution, superimposed coding.**

I. Introduction

Chemical structure retrieval systems are frequently presented with the task to produce a list of all stored chemical graphs containing a prescribed subgraph [1], [2]. Due to the absence of a linear order among the stored data, tree-based search strategies fail, and a sequential search has to be performed. To accelerate this time-consuming process, the actual graph theoretical substructure match is preceded by *prescreening*: the entire database is matched against a library of simple but common *descriptors*, and the validity of descriptors is recorded in a bitstring for each stored structure. Suitable choices for descriptors are small chemical subgraphs containing only few vertices, graph diameters, ring sizes, or any other property that passes from subgraphs to supergraphs. When a query structure is submitted to the system, the descriptors are evaluated for this query structure resulting in a query bitstring. Only those stored structures are candidates for a match where each bit is turned on in all those positions where the query bits are turned on and will be subjected to the expensive graph theoretical matching algorithm.

For example, let us consider the compounds in Fig. 1. A chemist might ask for a list of all structures in our database containing 2-(cyclohexylmethyl)naphthalene, which is too complex to be one of the index descriptors. However, any matching structure must necessarily contain cyclohexane and naphthalene, and these might be indexed. We will produce an intermediate result set that also contains 2-(2-cyclohexylethyl)naphthalene, which is not in accordance with the original query specification and must be singled out by graph matching.

The Beilstein database of organic compounds contains around ten million structures, each a chemical graph of up to 255 vertices, and the number of descriptors will have the magnitude of one thousand. With such characteristics, the preevaluated bitstrings will consume a considerable amount of storage and hence of processing time. On the other hand, one may expect the 1 bits to be relatively scarce, whence it should be possible to compress the bitstrings without losing too much information. Thus, we are looking for a map $\psi : \dot{I}^N \to \dot{I}^n, \dot{I} = \{0, 1\}$ transforming bitstrings of length N into strings of length n. However, we have to observe the partial order relation $\beta \leq \beta'$ between bitstrings defined such that the bits should satisfy $\beta(i) \leq \beta'(i)$ at all positions i. Since we want to use the compressed strings in the same manner (by bitmasking) as the original ones, the transformation must be monotone: $\beta \leq \beta' \Rightarrow \psi(\beta) \leq \psi(\beta')$, and in particular,

Manuscript received April 12, 2007; revised February 18, 2008.

The author is with the Beilstein-Institut, Frankfurt 60487, Germany (e-mail: dr.bernd.guenther@t-online.de).

Communicated by V. A. Vaishampayan, Associate Editor At Large.

Digital Object Identifier 10.1109/TIT.2008.924658