ABSTRACT

Title of dissertation: PERFORMANCE ANALYSIS OF
ALGEBRAIC SOFT-DECISION DECODING
OF REED-SOLOMON CODES

Andrew Duggan, Masters of Science, 2006

Thesis directed by: Professor Alexander Barg
Department of Electrical and Computer Engineering

We investigate the decoding region for Algebraic Soft-Decision Decoding (ASD) of Reed-Solomon codes in a discrete, memoryless, additive-noise channel. An expression is derived for the error radius within which the soft-decision decoder produces a list that contains the transmitted codeword. The error radius for ASD is shown to be larger than that of Guruswami-Sudan hard-decision decoding for a subset of low-rate codes. We then present an upper bound for ASD's probability of error, where an error is defined as the event that the decoder selects an erroneous codeword from its list. This new definition gives a more accurate bound on the probability of error of ASD. We also derive an estimate of the error-correction radius under multivariate interpolation decoding of a recent generalization of Reed-Solomon codes by F. Parvaresh and A. Vardy.

PERFORMANCE ANALYSIS OF ALGEBRAIC SOFT-DECISION DECODING OF REED-SOLOMON CODES

by

Andrew Duggan

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Masters of Science
2006

Examination Committee:
Professor Alexander Barg, Chair/Advisor
Professor Adrian Papamarcou
Professor Sennur Ulukus

ACKNOWLEDGMENTS

My deep thanks goes to my advisor Professor Alexander Barg for his guidance and countless number of helpful conversations. This thesis is a product of our collaboration and relentless exploration of coding theory. By adopting his attention to detail, I learned the importance of composing a technical document that is both concise and precise.

I would also like to express my appreciation to the other members of the thesis evaluation committee: Professor Adrian Papamarcou, Professor Sennur Ulukus, and Paul Conoval. Their evaluation of my work helped give me added insight that is a product of a different perspective.

Thank you as well to my wife Alice Duggan for her support throughout my research process. It not always easy to live with somebody who experiences the highs and lows of critical thinking, but she provided a solid foundation throughout my career as a Masters student at the University of Maryland.

CONTENTS

List of Figures

1. INTRODUCTION

Error correcting codes were invented to improve communication across noisy channels. In his seminal 1948 paper [17], Shannon stated that there exists error correcting codes that can achieve an arbitrarily small probability of error given the rate of such a code falls below the channel's capacity. Since Shannon's claim, researchers have focused their attention on finding such code constructions, resulting in the invention of a multitude of code families.

Of all the code families in existence, none has likely been applied to such a wide array of real-world, engineering problems as Reed-Solomon (RS) codes. RS codes are being used to correct errors occurring in storage devices such as CDs, DVDs, and hard drives. One will often find RS codes on telecommunication lines such as satellite links. Indeed, advances in the decoding of Reed-Solomon codes would certainly find application in today's engineering problems.

2. LIST DECODING

In this chapter, we describe the Algebraic Soft-Decision Decoding (ASD) algorithm of RS codes and prove its correctness. The results presented here are mostly due to Guruswami and Sudan [7] with contributions by Koetter and Vardy [11] and McEliece [12]. We also discuss how ASD can be used to decode a family of RS-based codes recently proposed by Parvaresh and Vardy [13].

2.1. **Introduction.**

RS codes are one of the most studied class of error-correcting codes. Under the traditional approach [1], they are decoded to correct errors up to half of their minimum distance d. In 1999, Guruswami and Sudan [7] suggested a list-decoding algorithm of RS that corrects more than $^d/_2$ errors in the list-decoding sense. An even more powerful version of their algorithm, the Algebraic Soft-Decision Decoder, was introduced in the same paper [7] and later refined by Koetter and Vardy in [11]. While this approach shows promise to improve the decoding performance of Reed-Solomon codes, the improvement claim over conventional hard-decision decoding techniques in the published literature is only supported by experimentation.

This thesis gives an error radius within which ASD has the transmitted codeword on its list for a discrete, memoryless channel with additive noise. Based on a comparison of this radius with that of hard-decision decoding methods, we are indeed able to claim that ASD gives a performance improvement in RS decoding for a subset of low-rate codes. However, one should note that there exists decoder configurations which do not exhibit an improvement.

We present new results in upper bounding the probability of error for ASD. In [10], Ratnakar and Koetter derive an upper bound for the probability that the decoder's list does not contain the transmitted codeword. We show that the probability of error for this error event can be zero for low-rate codes, making it is not a comprehensive measure of ASD's probability of error. We redefine the definition of ASD's probability of error to be the selection of the correct codeword from the decoder's list, and we derive an upper bound for the probability for this new error event for ASD.

2.2. **Reed-Solomon Coding Model.**

Let q be a prime power and let $\mathbb{F}_q = \{\alpha_0 = 0, \alpha_1, ...\alpha_n\}$ be the finite field of q elements. For a polynomial $f \in \mathbb{F}[X]$ define the evaluation mapping eval : $f \to \mathbb{F}_q^n$ given by $(\text{eval}f)_i = f(\alpha_i)$, $1 \le i \le n$. Thus, the evaluation mapping associates a q-ary n-vector to every polynomial $f \in \mathbb{F}_q[x]$.

Definition 1. *A q-ary RS code C of length $n = q - 1$ and dimension k is the set of codewords of the form $\underline{c} = \text{eval}(f)$ where f runs over all polynomials over $\mathbb{F}_q$ of degree $0 \le \deg f \le k - 1$.*

To describe the encoding of the code C, suppose that the message to be transmitted is $\underline{u} = (u_1, u_2, ...u_k)$ where $u_i \in \mathbb{F}_q$, $1 \le i \le k$. The codeword that corresponds to it is given by $\underline{c} = \text{eval}(f)$, where the polynomial f has the form

$$f(X) = u_1 + u_2X + u_3X^2 + ... + u_kX^{k-1}.$$

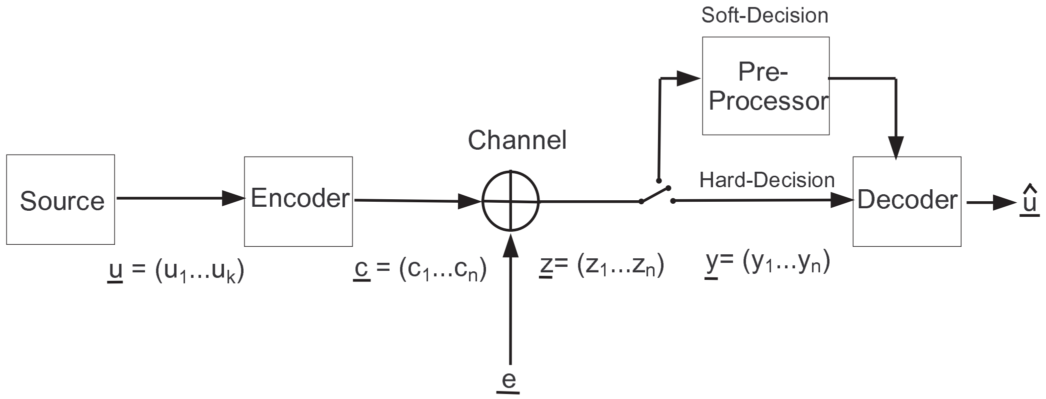
Soft-Decision
Pre-Processor
Channel
Source
Encoder
Hard-Decision
Decoder
$\hat{\underline{u}}$
$\underline{u}$ = (u₁...uₖ)
$\underline{c}$ = (c₁...cₙ)
$\underline{z}$= (z₁...zₙ)
$\underline{y}$= (y₁...yₙ)
$\underline{e}$

FIGURE 1. Block diagram of a communication system using an error-correcting code.

We assume that the codeword $\underline{c}$ is transmitted over a discrete, memoryless channel with additive noise $\underline{e}$ as shown in Figure 2.1. The output of the channel is the vector $\underline{z} = \underline{c} + \underline{e}$, where the alphabet for $\underline{z} = (z_1, z_2, ...z_n)$ is $\mathscr{Z} = \{\zeta_1, \zeta_2, ...\zeta_J\}$ that is not necessarily equal to $\mathbb{F}_q$. Define the channel transition probability as

$$w_{i,j} = \Pr(z = \zeta_i | c = \alpha_j).$$

Since the channel is additive and memoryless, we know it is symmetric as defined in Section 8.2 of [2].

RS codes can be decoded by either making hard decisions or soft decisions based on information gathered from the channel. If soft-decision decoding is utilized, then the channel's output is taken in the form of a q x n matrix Π defined as

$$\pi_{i,j} = \Pr(c = \alpha_i | z = z_j).$$

Each column of the matrix Π is the set of posterior probabilities for one symbol position of the codeword. As is explained below, the ASD algorithm takes as inputs nonnegative integers rather than probabilities. For this reason, a Pre-Processor converts the matrix Π into a $q \times n$ matrix $\mathcal{M}$ of non-negative integers that is passed to the decoder.

If hard-decision decoding is used, the vector $\underline{z}$ is converted to the vector $\underline{y} = (y_1, y_2, ...y_n)$, where

$$y_i = \text{argmax}_{x \in \mathbb{F}_q} \Pi(x, i). \tag{1}$$

The decoder then operates directly on $\underline{y}$ to give its best estimate of the message $\hat{\underline{u}}$.

2.3. **Reed-Solomon Decoding Techniques.** This thesis concentrates on the ASD algorithm [11] and compares its performance to the well-known hard-decision decoding algorithms of Berlekamp-Massey (see e.g. [1]) and Guruswami-Sudan [7, 12]. Most decoders in use currently use some variant of Berlekamp-Massey (BM) syndrome decoding.

Let $\underline{y}$ be the hard-decision vector formed according to (1), and let $\underline{c}$ be the transmitted codeword. Let dist(·,·) be the Hamming distance. If the number of errors $t = \text{dist}(\underline{c}, \underline{y})$ satisfies

$$t \leq \left\lfloor \frac{n-k}{2} \right\rfloor, \tag{2}$$

then the decoder will output $\underline{c}$. If condition (2) is not true, then decoding is guaranteed to fail. Therefore, (2) is a necessary condition for BM decoding success.

Guruswami-Sudan (GS) decoding produces a list that contains all the codewords of distance t_m from the vector $\underline{y}$ and potentially some codewords outside of this Hamming ball. List decoding success is declared if the correct codeword is on the list. Figure 2.2 is a conceptual picture of GS decoding in the Hamming space over the field $\mathbb{F}_q^n$. In this case, $\underline{c}_1$ was transmitted, $\underline{y}$ was received, and the channel has caused t errors. With

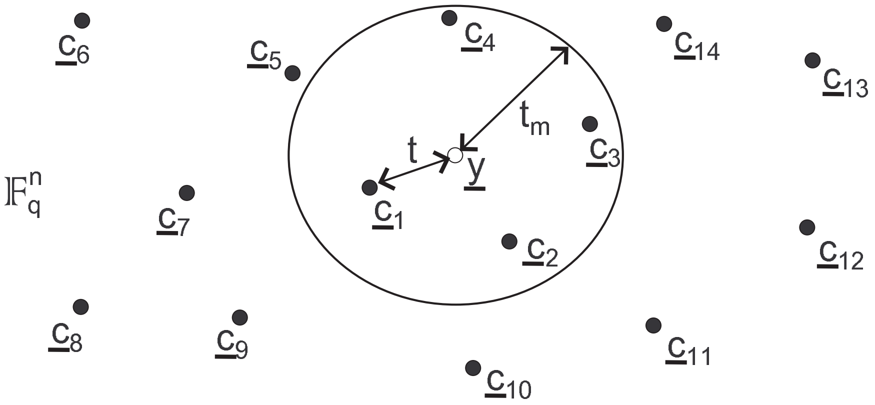

FIGURE 2. GS Decoding in Hamming space.

the decoding distance at t_m, GS will produce a list of at least four codewords that contains the transmitted codeword. In the list decoding sense, Figure 2.2 is a picture of successful decoding.

The distance t_m is determined by m which is a parameter of the algorithm. As m increases, t_m increases to an asymptotic limit given in Lemma 1.

Lemma 1. (Guruswami and Sudan [7]) *Let $m \to \infty$. Let $\underline{c}$ be a codeword that satisfies*

$$\text{dist}(\underline{y}, \underline{c}) < n - \sqrt{nk}. \tag{3}$$

Then $\underline{c}$ will be included in the list output by the GS decoder with input $\underline{y}$.

The complexity of the algorithm often becomes a limiting factor before the maximum possible t_m is achieved. Note that GS decoding is guaranteed to have the transmitted codeword on the list if the error pattern satisfies (3), assuming large m. However, it is not necessarily true that an error pattern that does not satisfy (3) will not be on the list. Thus, (3) is only a sufficient condition on GS list-decoding success.

Let $\tau = t/n$ be the normalized error correction radius of RS decoding algorithms, and let $R = k/n$ be the rate of the code. We have

$$\tau = \frac{1-R}{2} \text{ (BM Decoding)} \qquad \tau = 1 - \sqrt{R} \text{ (GS Decoding, } m \text{ large).} \tag{4}$$

Figure 2.3 compares the error radii given by (4). The GS decoding radius is always greater than its BM counterpart as shown in Figure 2.3; however, the difference becomes small for high rates. Missing from the graph and from the published literature is a sufficient condition for ASD list-decoding success.

2.4. **Bivariate Polynomials.** In this section, we prepare the way for the description of algebraic list decoding algorithms. Bivariate polynomials play a central role in ASD and GS decoding, so it is necessary first to introduce some basic concepts. Let

$$Q(X,Y) = \sum_{i,j} a_{i,j} X^i Y^j \tag{5}$$

be a polynomial over $\mathbb{F}_q$. For a given positive integer v, define the weighted degree of the monomial X^iY^j as

$$\text{wdeg}_v\, X^iY^j = i + vj.$$

We will often use the value $v = k - 1$ for the weight and omit the subscript v in this case. For v = 1, the weighted degree will be called simply the degree. The weighted degree of a bivariate polynomial $Q(X,Y)$ is equal to maximum weighted degree among its component monomials. The weighted degree defines a reverse lexicographic order of monomials. Namely, let $X^{i_1}Y^{j_1}$ and $X^{i_2}Y^{j_2}$ be monomials with weighted

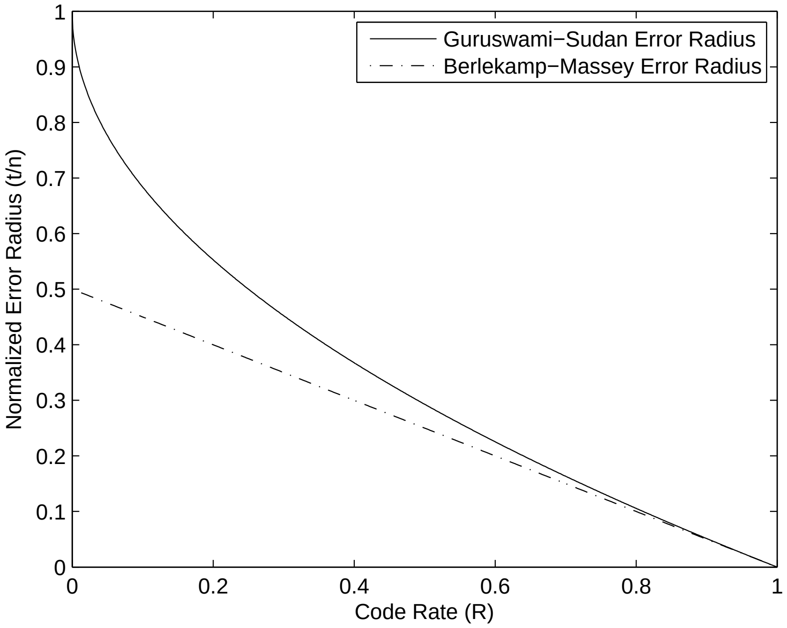

FIGURE 3. Performance of BM and GS Decoding.

degrees $w_1 = i_1 + vj_1$ and $w_2 = i_2 + vj_2$, respectively. We say that $X^{i_1}Y^{j_1} \prec X^{i_2}Y^{j_2}$ if $w_1 < w_2$ or if $i_1 < i_2$ in the case that $w_1 = w_2$. Furthermore, let

$$N_v(\delta) = |\{X^iY^j : i, j \geq 0 \;\; \& \;\; i + jv \leq \delta\}|.$$

It is of interest to determine the index of the term Y^L, denoted $B(L, v)$, and the index of the term X^K, denoted $A(K, v)$, within the reverse lexicographic order of monomials. Lemma 2 gives an expression for both of these quantities.

Lemma 2. (McEliece [12]) *Let* $r = K \; mod \; v$. *Then*

$$A(K, v) = \frac{K^2}{2v} + \frac{K}{2} + \frac{r(v-r)}{2v}$$

$$B(L, v) = \frac{vL^2}{2} + \frac{(v+2)L}{2}.$$

Lemma 2 implies

Corollary 1. *For* $v \geq 1$,

$$L = \left\lfloor \sqrt{\frac{2B(L,v)}{v} + \left(\frac{v+2}{2v}\right)^2} - \left(\frac{v+2}{2v}\right) \right\rfloor$$

$$K < \sqrt{2vA(k,v)}.$$

Finally, let us discuss what it means for $Q(X, Y)$ to pass through a point $(\alpha, \beta) \in \mathbb{F}_q^2$. The polynomial $Q(X, Y)$ has a zero at the point (α, β) (passes through the point (α, β)) if $Q(X - \alpha, Y - \beta) \equiv 0$. A polynomial can also pass through the same point many times, meaning that it has a multiple zero at that point.

Definition 2. *A bivariate polynomial $Q(X,Y)$ has a zero of multiplicity of order m at (0,0) if no monomial of $Q(X,Y)$ is of degree less than m. Similarly, $Q(X,Y)$ has a zero of multiplicity m at (α,β)$\in$ $\mathbb{F}_q^2$ if $Q(X+\alpha,Y+\beta)$ has a zero of multiplicity m at $(0,0)$.*

For example, the polynomial

$$Q(X,Y)=(X-\alpha)^2(Y-\beta)+(X-\alpha)(Y-\beta)^2$$

has a zero of multiplicity 3 at (α,β).

2.5. **Guruswami-Sudan List Decoding.** Intuitively, the GS decoder determines its list by curve-fitting all the codewords to the received vector $\underline{y}$. Consider the mapping mult: $\mathbb{F}_q^n \to \{(\alpha_1,\beta_1),...(\alpha_n,\beta_n)\}$ defined by (mult $\underline{y}) = \{(x_1,y_1),...(x_n,y_n)\}$. A GS decoder attempts to find codewords that are a good fit to the set of points (mult $\underline{y}$). The codewords that match in at least t_m symbol positions are included in the list by the decoder. GS decoding consists of the three steps detailed in Sections 2.5.1 and 2.5.2.

2.5.1. *Interpolation.* The first phase of decoding is called interpolation since it consists of filling in the space between points with a curve. When the decoder receives $\underline{y}$, it constructs a bivariate polynomial $Q(X,Y)$ that has a zero of multiplicity m at each of the points (x_i,y_i). This task is accomplished by solving the system of linear equations

$$Q(X+x_i,Y+y_i)\equiv 0, \quad i=0,1,\dots,n. \tag{6}$$

For each i, the decoder has to find a $Q(X+x_i,Y+y_i)$ that has a zero of multiplicity m at $(0,0)$. This implies that all the monomial coefficients of $Q(X+x_i,Y+y_i)$ that are of degree less than m must be zero. This task is difficult with the system linear equation expressed as (6). The following lemma allows us to rewrite these equations into an easier form to handle.

Lemma 3. (Guruswami and Sudan [7]) *Define $Q(X,Y)$ as in (5). For any $(\alpha,\beta)\in\mathbb{F}_q^2$,*

$$Q(X+\alpha,Y+\beta)=\sum_{r,s}Q_{r,s}(\alpha,\beta)X^rY^s,$$

where

$$Q_{r,s}(X,Y)=\sum_{r,s}\binom{i}{r}\binom{j}{s}a_{i,j}X^{i-r}Y^{j-s}.$$

Proof.

$$\begin{aligned}Q(X+\alpha,Y+\beta)&=\sum_{i,j}a_{i,j}(X+\alpha)^i(Y+\beta)^j\\&=\sum_{i,j}a_{i,j}\Big(\sum_r\binom{i}{r}X^r\alpha^{i-r}\Big)\Big(\sum_s\binom{j}{s}Y^s\alpha^{j-s}\Big)\\&=\sum_{r,s}X^rY^s\Big(\sum_{i,j}\binom{i}{r}\binom{j}{s}a_{i,j}\alpha^{i-r}\beta^{j-s}\Big)\\&=\sum_{r,s}Q_{r,s}(\alpha,\beta)X^rY^s.\end{aligned}$$

□

The expression $Q_{r,s}(X,Y)$ is called the (r,s)th Hasse derivative of $Q(X,Y)$. Lemma 3 allows us to rewrite (6) as

$$Q_{r,s}(x_i,y_i)=0,\ \ \forall i:0\le i<n,\ \ \forall(r,s):0\le r+s<m. \tag{7}$$

Methods to solve the set of linear equation (7) efficiently have been suggested by Feng-Tzeng [4] and Koetter [10]; Gaussian elimination could also be used to solve (7) but not as efficiently as the two aforementioned methods. Lemma 4 gives a condition under which there exists a non-zero polynomial $Q(X, Y)$ that satisfies (7).

Lemma 4. *Let* wdeg $Q(X, Y) = \delta$. *If*

$$N_{k-1}(\delta) > n\binom{m+1}{2},$$

then the system (7) has a non-zero solution.

Proof. Let us determine the number of linear equations contained in (7). For each i, the number of equations equals the pairs of integers that satisfies $0 \leq r + s < m$. This number is $\frac{m(m+1)}{2}$, and the total number of equations (constraints) is $n\frac{m(m+1)}{2}$.

The number of unknowns is the number of coefficients of $Q(X, Y)$, or $N_{k-1}(\delta)$. Thus, if $N_{k-1}(\delta) > n\binom{m+1}{2}$, the number of unknowns exceeds the number of constraints, and there must be a non-zero solution. □

2.5.2. *Factorization and Selection.* It is not required to fully factor $Q(X, Y)$ in order to determine the decoder's list. We only need to determine the factors of $Q(X, Y)$ of the form $Y - f(X)$ where $\deg f(X) \leq k - 1$. An efficient root-finding technique has been suggested by Roth and Ruckenstein in [16].

A key goal in list decoding is bounding the size of the output list. A simple upper bound can be derived by finding an expression for the Y-degree of $Q(X, Y)$. This bound will not be tight due to the fact that $Q(X, Y)$ may have roots that satisfy (6) but do not correspond to codewords because their degree exceeds $k - 1$. This upper bound on the size of the list can be found by considering the Y^L monomial that has the same degree as greatest Y-degree monomial of $Q(X, Y)$. Lemma 5 gives an upper bound for the value of L.

Lemma 5. (McEliece [12]) *The size of the list L produced by a GS decoder is bounded above as*

$$L < \left\lfloor \sqrt{\frac{n}{k-1}m(m+1) + \left(\frac{k+1}{2k-2}\right)^2} - \left(\frac{k+1}{2k-2}\right) \right\rfloor. \tag{8}$$

Proof. Assume that Y^L has the same Y-degree as $Q(X, Y)$; it follows that L bounds the size of the list produced by a GS decoder. Corollary 1 gives an upper bound on L, so it is only a matter of determining $B(L, k - 1)$ and v. We set $v = k - 1$ because we are interested in polynomials of degree $k - 1$ or less. Since Y^L is the largest monomial of weighted degree L, it follows that $B(L, k - 1) = N_{k-1}(L)$. By using Lemma 4, we know $Q(X, Y)$ exists if

$$B(L, k-1) > n\binom{m+1}{2}.$$

Setting $B(L, k - 1) = n\binom{m+1}{2}$ gives us the upper bound in the lemma. □

After factorization is complete, maximum likelihood selection for the list obtained can be is used to determine the best estimate of the transmitted message. In this case, maximum likelihood selection consists of a search for the codeword on the list that is closest to the received vector by Hamming distance.

2.6. **Algebraic Soft-Decision Decoding.** ASD extends GS decoding through the manipulation of the multiplicities. Instead of operating on a vector $\underline{y}$, ASD takes as input multiplicity matrix $\mathcal{M}$. The soft-decision decoder constructs a bivariate polynomial $Q(X, Y)$ that has zeros of multiplicity set by $\mathcal{M}$. In contrast to GS decoding that always constructs $Q(X, Y)$ based on n distinct zeros, ASD can have up to qn distinct zeros. The methods of determining $\mathcal{M}$ from the channel output is discussed in Section 2.6.1.

Similarly to GS decoding, ASD uses curve-fitting of codeword polynomials to points. However, since there are many more points, including multiple points for the same codeword position, Hamming distance is an inadequate way to measure the fit of a codeword. Thus, we must introduce the notion of a codeword score.

Definition 3. *Let the score of a vector* $\underline{v}$ *be*

$$\langle \mathcal{M}, [\underline{v}] \rangle = \sum_{i=1}^{n} m_{v(i),i}.$$

The notation $[.]$ *represents the indicator matrix such that* $[\underline{v}]_{i,j} = 1$ *if* $v_j = \alpha_i$*, where* $\mathbb{F}_q = \{\alpha_1, \alpha_2, ...\alpha_q\}$*, and* $[\underline{v}]_{i,j} = 0$ *otherwise. Furthermore,* $v(i) = l$ *where* $v_i = \alpha_l$*.*

The higher the score of a codeword, the better the fit to the multiplicity matrix. In this new paradigm, the geometric picture of Figure 2.2 no longer applies. The question that follows is how we can now determine whether a codeword is on the list produced by ASD. This question is answered in Section 2.6.2.

2.6.1. *The Multiplicity Matrix.* The matrix $\mathcal{M}$ is determined from the matrix of posterior probabilities Π which is based on the received vector $\underline{z}$. This vector $\underline{z}$ is equal to the hard-decision vector $\underline{y}$ only if $\underline{e} \in \mathbb{F}_q^n$. Koetter-Vardy [11], Parvaresh-Vardy [13], and El-Khamy-McEliece [3] have proposed various methods for determining $\mathcal{M}$ from Π. This thesis will use the simplest method for converting Π to $\mathcal{M}$ proposed by Gross-Kschischang-Koetter-Gulak [6]. It is named the Proportionality Multiplicity Assignment Strategy (PMAS), and it finds $\mathcal{M}$ by performing the following element-wise calculation on Π.

$$m_{i,j} = \lfloor \lambda \pi_{i,j} \rfloor \tag{9}$$

The parameter $\lambda \in \mathbb{Z}^+$ is the complexity factor, and its adjustment directly controls the balance between the performance and the complexity of ASD. Another important measure of the complexity of the decoder is the cost of the multiplicity matrix.

Definition 4. *Let the Cost of a multiplicity matrix* $\mathcal{M}$ *be*

$$\mathcal{C}(\mathcal{M}) = \frac{\langle \mathcal{M}, \mathcal{M} \rangle + \langle \mathcal{M}, 1 \rangle}{2} = \frac{1}{2} \sum_{i,j} m_{i,j}(m_{i,j} + 1).$$

2.6.2. *Threshold Condition.* Since we are no longer guaranteed to produce a list that contains all the codewords less than Hamming distance t_m, one wonders how to know if a codeword $\underline{c}$ is on the list produced by ASD. It turns out that if $S_{\mathcal{M}}(\underline{c})$ exceeds a threshold given in Lemma 8, it follows that $\underline{c}$ is on the list produced by ASD.

In order to motivate Lemma 8, consider the following system of linear equations associated with a given multiplicity matrix $\mathcal{M}$.

$$Q_{r,s}(x_i, y_i) = 0, \;\; \forall (i,j,r,s) : 0 \le r + s < m_{i,j} \tag{10}$$

Claim 1. *The number of equations in (10) equals* $\mathcal{C}(\mathcal{M})$*.*

Similarly to Lemma 4, a polynomial $Q(X,Y)$ that satisfies (10) is guaranteed to exist if

$$N_{k-1}(\text{wdeg } Q(X,Y)) > \mathcal{C}(\mathcal{M}).$$

We need to introduce one technical lemma before proceeding.

Lemma 6. (Guruswami and Sudan [7]) *If* (α, β) *is a zero of multiplicity* m *of polynomial* $Q(X,Y)$ *and* $\beta = f(\alpha)$*, then* $(X - \alpha)^m \mid Q(X, f(X))$*.*

Lemma 7 will show that $Y - f(X)$ is a factor of such a $Q(X,Y)$ when the score of the codeword is large enough.

Lemma 7. *Let* $\delta = \text{wdeg } Q(X,Y)$ *and let* $\underline{c} = \text{eval}(f)$ *where* $f \in \mathbb{F}_q[x]$ *is a polynomial of degree* $\le k-1$*. If* $N_{k-1}(\delta) > \mathcal{C}(\mathcal{M})$ *and* $S_{\mathcal{M}}(\underline{c}) > \delta$*, then* $(Y - f(X)) \mid Q(X,Y)$*.*

Proof. Define the polynomial $P(X)$ as

$$P(X) = Q(X, f(X)).$$

If it can be shown that $P(X) \equiv 0$, then $Y - f(X)$ is a factor of $Q(X, Y)$. Let $m_i = m_{c(i),i}$ and $S_{\mathcal{M}}(\underline{c}) = \sum_{i=1}^{n} m_i$, i.e. $Q(X, Y)$ has zeroes of multiplicity m_i at each of the points (x_i, c_i), respectively. From Lemma 6, we have

$$(X - x_1)^{m_1}(X - x_2)^{m_2}...(X - x_n)^{m_n} \mid P(X).$$

We now know that either deg $P(X) \geq S_{\mathcal{M}}(\underline{c})$ or $P(X) \equiv 0$. On the other hand, since deg $f(X) \leq k - 1$,

$$\text{deg } P(X) \leq \text{wdeg } Q(X, Y).$$

Therefore, if $S_{\mathcal{M}}(\underline{c}) >$ wdeg $Q(X, Y)$, then it follows that $P(X) \equiv 0$, and the lemma is proven. □

Lemma 8 is the main result of the section. It follows directly from Lemma 6.

Lemma 8. (Koetter and Vardy [11]) *If*

$$S_{\mathcal{M}}(\underline{c}) > \sqrt{2(k-1)\mathcal{C}(\mathcal{M})}$$

or equivalently

$$\sum_{i=1}^{n} m_{c(i),i} > \sqrt{(k-1)\sum_{i,j} m_{i,j}(m_{i,j}+1)},$$

then $Q(X, Y)$ contains a factor $Y - f(X)$.

Proof. If we can find an upper bound for wdeg $Q(X, Y)$, and if the score of a codeword exceeds this upper bound, then we know from Lemma 7 that this codeword is on the decoder's list. To this end, consider the monomial X^K that is same degree as the greatest weighted degree among the monomials of $Q(X, Y)$. Corollary 1 provides an upper bound on K as follows.

$$K < \sqrt{2vA(K, v)}$$

We set $v = k - 1$ since we trying to find a factors of the form $Y - f(X)$ where deg $f(X) \leq k - 1$. We also know that $A(K, v) \geq \mathcal{C}(\mathcal{M})$, otherwise there may not exist a $Q(X, Y)$ which meets all the linear constraints of $\mathcal{M}$. Thus, $\sqrt{2(k-1)\mathcal{C}(\mathcal{M})}$ is an upper bound for wdeg $Q(X, Y)$. If $S_{\mathcal{M}}(\underline{c})$ exceeds this upper bound, then the conditions of Lemma 7 must be true, and Lemma 8 is proven. □

2.7. **Multivariate Interpolation.** Recently, Parvaresh and Vardy introduced a new class of codes constructed as evaluations of $M \geq 2$ polynomials and extended the decoding procedure described above to multivariate interpolation. A code C in the Parvaresh-Vardy (PV) family is defined as follows.

Let $\{1, \beta_1, \beta_2, ...\beta_{M-1}\}$ be a basis over $\mathbb{F}_{q^M}$, let $\{a_1, a_2, ...a_{M-1}\}$ be a set of positive integers greater than 1, and let $e(X)$ be an irreducible polynomial over $\mathbb{F}_q$. The code has parameters $[n, k]$ where $n = q - 1$. It follows that the rate of a PV code C is $R = \frac{k}{Mn}$ and the minimum distance is $d = n - k + 1$. The encoder constructs $f(X)$ as the polynomial derived from the message $\underline{u}$, and it finds the set of polynomials $\{g_1(X), g_2(X), ...g_{M-1}(X)\}$ by computing

$$g_i(X) = (f(X))^{a_i} \mod e(X). \tag{11}$$

Since (11) is a non-linear operation, the code C is not necessarily linear. A codeword $\underline{c} = \{c_1, c_2, ...c_n\}$ of a PV code that is associated with $\underline{u}$ is found through the evaluation

$$c_i = f(\alpha_i) + \sum_{j=1}^{M-1} \beta_j g_j(\alpha_i), \quad \forall i : 1 \leq i \leq n,$$

where$\alpha_i, i = 1, \ldots, n$ are all the nonzero elements of $\mathbb{F}_q$.

The authors considered only hard decision decoding by an extension of the Guruswami-Sudan method; however, following Lemmas 6-8, it is possible to establish a sufficient condition under which soft-decision decoding places a codeword on the list.

Theorem 1. *Let C be an $[n,k]$ PV code over $\mathbb{F}_{q^M}$ communicated over a discrete memoryless channel with additive noise. Suppose that it is decoded using a multivariate version of the ASD algorithm. A codeword $c = (c_1, \ldots, c_n)$ will be included in the list output by the algorithm if*

$$\sum_{i=1}^{n} m_{c(i),i} > \sqrt[M+1]{(k-1)^M \sum_{i,j} \binom{m_{i,j}+M}{m_{i,j}-1}}.$$

Proof. By extending equation (38) of [13], we can derive an upper bound for the multivariate polynomial as

$$\text{wdeg } Q(X, Y_1, ...Y_M) < \left\lfloor \sqrt[M+1]{(k-1)^M \sum_{i,j} \prod_{l=0}^{M} (m_{i,j}+l)} \right\rfloor. \tag{12}$$

where wdeg $X^i Y_1^{j_1} ... Y_M^{j_M} = i + (k-1)\sum_{l=1}^{M} j_l$. If the score $S_{\mathcal{M}}(\underline{c})$ exceeds (12), then $\underline{c}$ is on the algebraic soft-decision decoder's list by a similar argument that was used in Lemma 7. □

3. ASD Error Correcting Performance

This chapter quantifies the performance of Algebraic Soft-Decision Decoding in two ways. First, we derive an ASD decoding radius and show that this radius is larger than the GS decoding radius of Lemma 1 for a subset of low-rate codes for discrete, memoryless channels with additive noise. The only exception to this claim is if the complexity factor does not exceed a threshold that is a function of the channel.

Second, we consider the ASD probability of error by realizing that decoding success is the intersection of the decoder's list containing the transmitted codeword and the decoder selecting the correct codeword from the list. While the published literature has taken the list-decoding probability of error to only measure the first event, we propose that both events must be considered to give good insight into the list-decoding probability of error. An upper bound is presented for this newly-defined probability of error termed the probability of error for decoder selection.

3.1. Setting.

Assume that the error vector $\underline{e} \in \mathbb{F}_q^n$. Thus, $\underline{z} = \underline{y}$. We also assume that each symbol entering the channel is uniformly drawn from $\mathbb{F}_q$, an assumption also made by Koetter and Vardy in [11] and Justesen in [9]. It follows that $\pi_{i,j} = w_{i,j}$. Since the channel is symmetric, we can assume that the channel transition probabilities $w_{i,j}$ are drawn from the set $\{p_1, p_2, ...p_q\}$. Next, we will introduce the three channel statistics

$$p_{\max} = \max_{1\le i\le q} p_i \qquad p_{\min} = \min_{\substack{1\le i\le q\\ i:p_i>0}} p_i \qquad \gamma = \sum_{i=1}^{q} p_i^2.$$

When the channel is noiseless, set $p_{\min} = 0$. We will assume throughout that the channel's capacity is greater than zero, giving us $p_{\max} > p_{\min}$. As will be seen later, $p_{\max}$, $p_{\min}$, and γ will be the only channel statistics necessary in our analysis of ASD's performance.

3.2. ASD Error Radius.

In this section, we present one of our main results, an estimate of the error correction radius of the algorithm. ASD will produce a list with a codeword $\underline{c}$ when $\underline{y}$, as defined in (1), is within a Hamming distance t from $\underline{c}$. Thus, we are analyzing a soft-decision decoder using a hard-decision metric. We would like to stress that the error correction radius of ASD decoding is defined in a different way than for BM and GS decoding. In the case of BM and GS decoding, all of the codewords within the error radius are on the decoder's list, but for ASD, the only codeword within the error radius that is guaranteed to be on the list is the transmitted codeword.

Theorem 2. *Suppose a RS code with rate $R = k/n$ is used to communicate over an discrete, additive-noise channel. An algebraic soft-decision decoder, with complexity factor λ, is used to decode, and $t = \mathrm{dist}(\underline{c}, \underline{y})$. If*

$$\frac{t}{n} \leq \frac{p_{\max} - \sqrt{R\left(\gamma + \frac{1}{\lambda}\right)} - \frac{1}{\lambda}}{p_{\max} - p_{\min}}, \tag{13}$$

then the decoder's list will contain the transmitted codeword $\underline{c}$.

Proof. Let $\underline{c}$ be the transmitted codeword, let $\underline{y}$ be defined as in (1), and let the functions $c(i)$ and $y(i)$ be defined by $\alpha_{c(i)} = c_i$ and $\alpha_{y(i)} = y_i$, respectively. Substituting (9) in Lemma 8, we get

$$\frac{\sum_{i=1}^{n} \lfloor \lambda w_{c(i),y(i)} \rfloor}{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{q} \left(\lfloor \lambda w_{i,j} \rfloor^2 + \lfloor \lambda w_{i,j} \rfloor \right)}} \geq \sqrt{k-1}. \tag{14}$$

Introduce the parameter $\mu_{i,j}$ as the PMAS error term satisfying $m_{i,j} = \lambda \pi_{i,j} - \mu_{i,j} \quad \forall (i,j)$. The inequality (14) becomes

$$\frac{\sum_{i=1}^{n} \left(\lambda w_{c(i),y(i)} - \mu_{c(i),y(i)} \right)}{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{q} \left((\lambda w_{i,j} - \mu_{i,j})^2 + \lambda w_{i,j} - \mu_{i,j} \right)}} \geq \sqrt{k-1}. \tag{15}$$

Dropping the PMAS error terms in the denominator of the LHS of (15) can only make the LHS smaller. A smaller LHS makes (15) a more stringent condition for $\underline{c} \in \mathcal{L}$, where $\mathcal{L}$ is the soft-decision decoder's list. Rearranging and using the additive nature of the channel yields

$$\frac{1}{n} \sum_{i=1}^{n} w_{c(i),y(i)} \geq \sqrt{R\left(\gamma + \frac{1}{\lambda}\right) - \frac{\gamma}{n} - \frac{1}{\lambda n}} + \frac{1}{\lambda n} \sum_{i=1}^{n} \mu_{c(i),y(i)}. \tag{16}$$

Inequality (16) certainly holds true for the codeword $\underline{c}$ if

$$\frac{1}{n} \sum_{i=1}^{n} w_{c(i),y(i)} \geq \sqrt{R\left(\gamma + \frac{1}{\lambda}\right)} + \frac{1}{\lambda}. \tag{17}$$

We have derived a condition based on a specific $\underline{y}$, but we are interested in ASD's performance for any $\underline{y}$ when $\underline{c}$ is transmitted. Thus, $\underline{y}$ becomes a random variable. Let $\mathbf{W}$ be $w_{c(i),y(i)}$ given random $\underline{y}$. Thus, $\mathbf{W}$ is also a random variable. The dependence on $\underline{c}$ has been dropped since the p.m.f. of $\mathbf{W}$ is independent of $\underline{c}$ for an additive channel. The p.m.f. of $\mathbf{W}$ is as follows.

$$p_W(p_i) = \Pr\{\mathbf{W} = p_i\} = p_i \; \forall i : 1 \leq i \leq q$$

Taking each $\mathbf{W}_i$ to be a random variable with the same p.m.f. as $\mathbf{W}$, equation (17) can be rewritten

$$\frac{1}{n} \sum_{i=1}^{n} \mathbf{W}_i \geq \sqrt{R\left(\gamma + \frac{1}{\lambda}\right)} + \frac{1}{\lambda}. \tag{18}$$

Using t as defined in the theorem statement, the LHS of (18) can be lower below as

$$\frac{1}{n} \sum_{i=1}^{n} \mathbf{W} \geq \frac{1}{n}(t p_{\min} + (n-t) p_{\max}).$$

Thus, $\underline{c}$ is on the soft-decision decoder's list if

$$\frac{1}{n}(t p_{\min} + (n-t) p_{\max}) \geq \sqrt{R\left(\gamma + \frac{1}{\lambda}\right)} + \frac{1}{\lambda}.$$

The theorem's result follows. □

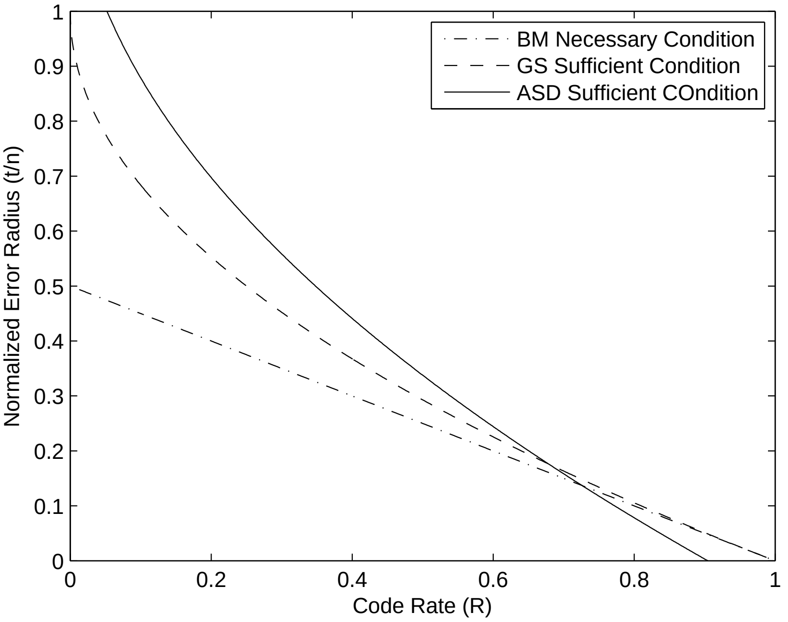

FIGURE 4. Decoding radius for the "typewriter channel" of Example 1.

A first look at the error radius in Theorem 2 reveals that (13) becomes large as $p_{\min}$ approaches $p_{\max}$. In other words, ASD performs well when the channel is far from q-ary symmetric. If the channel is noiseless and λ is sufficiently large, then the bound in Theorem 2 reduces to $1-\sqrt{R}$ which is the GS normalized error bound. Let us consider two examples.

Example 1. *Consider the "typewriter channel" where $w_{i,i}=0.8$, $w_{i,j}=0.2$ for some $j\neq i$, and $w_{i,j}=0$ for all the remaining pairs (i,j). Thus, $p_{\max}=0.8$, $p_{\min}=0.2$, and $\gamma=0.68$. Let us set $\lambda=100$. Figure 3.1 shows the normalized error bound compared to the GS error bound for this example. The BM error bound is also shown for reference. ASD is able to produce a list with the codeword $\underline{c}$ for a greater error radius than GS decoding for many low to medium rates. The range of rates for which ASD decoding corrects more errors than GS decoding is characterized in Section 3.4.*

Example 2. *Figure 3.2 shows the normalized error bound compared to the GS error bound for a q-ary symmetric channel with $p_{\max}=0.805$ and $q=16$. Thus, $p_{\min}=0.013$ and $\gamma=0.6506$, and we set $\lambda=\left\lceil\frac{1}{p_{\min}}\right\rceil=77$. ASD still provides an improvement, but it is only for extremely low-rate codes. The lack of improvement for the q-ary symmetric channel is expected since all error patterns, given that there are t errors, are equally probable.*

3.3. **Size of the List.** Proposition 1 gives an upper bound for the size of the list produced by ASD.

Proposition 1. *Given an additive noise channel, the size of the list for an algebraic soft-decision decoder, with $R>\frac{1}{n}$, is bounded above as*

$$|\mathcal{L}|\leq\left\lfloor\sqrt{\frac{\lambda^2\gamma+\lambda}{R-\frac{1}{n}}+\left(\frac{R+\frac{1}{n}}{2R-\frac{2}{n}}\right)^2}-\left(\frac{R+\frac{1}{n}}{2R-\frac{2}{n}}\right)\right\rfloor. \tag{19}$$

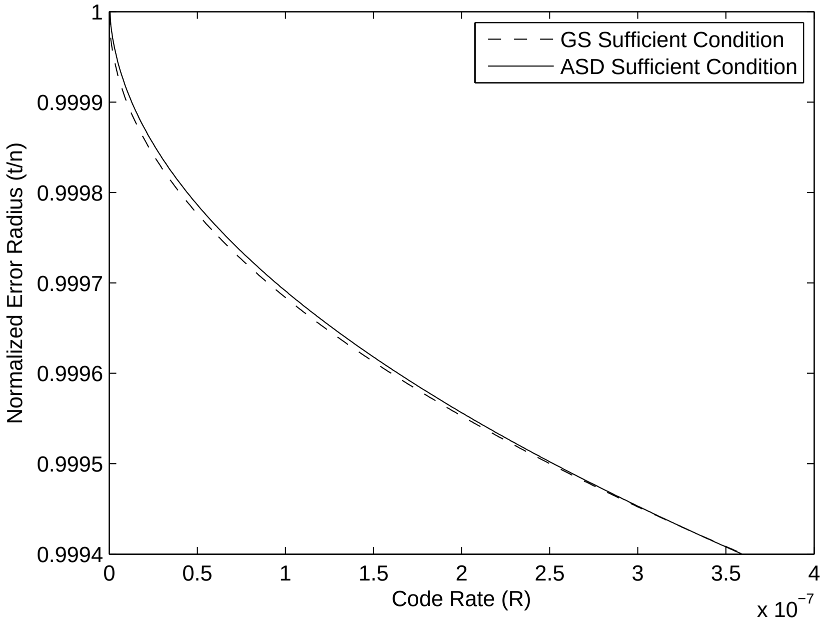

FIGURE 5. Decoding radius of ASD for a q-ary symmetric channel compared to GS and BM decoding.

Proof. Observe that

$$(20)\qquad \langle \mathcal{M},\mathcal{M}\rangle + \langle \mathcal{M},1\rangle = n\sum_{i=1}^{q}(\lambda p_i-\mu_i)^2 + n\sum_{i=1}^{q}(\lambda p_i-\mu_i)$$
$$\le n\lambda^2\gamma + n\lambda.$$

Substituting the upper-bound of (20) in (8) gives the result in the proposition. Since $v \geq 1$, it follows that $R > 1/n$. □

The bound in Theorem 1 is based on the Y-degree of the polynomial $Q(X,Y)$. Thus, it counts all such solutions to the system of equations (10), including the higher-order polynomials. As a result, Theorem 1's bound is not tight.

Example 3. *Figure 3.3 shows a graph of the bound on the list size presented in Theorem 1 for Example 1 with $n = 255$. The list size can be seen to be a strictly decreasing function of the rate.*

The maximum number of codewords on a list occurs when R is at its minimum. With the restriction $R > 1/n$ in place and the bound being a strictly decreasing function of the rate, an upper bound on the list size is obtained when we consider $R = 2/n$. In this case, (19) becomes

$$(21)\qquad |\mathcal{L}| \leq \left\lfloor \sqrt{n\lambda^2\gamma + n\lambda + \frac{9}{4}} - \frac{3}{2} \right\rfloor < \sqrt{\lambda^2\gamma+\lambda}\sqrt{n}.$$

As one can see from (21), the number of codewords is polynomial in n.

3.4. **A Closer Look at the ASD Error Radius.** We are interested in quantifying when the radius in Theorem 2 is larger than the radius in Lemma 1.

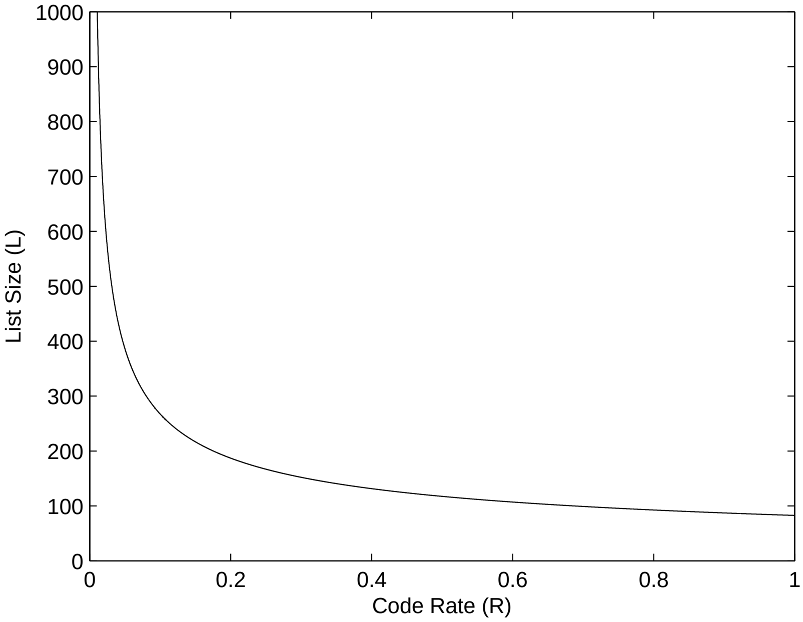

FIGURE 6. List size using ASD for the "typewriter channel".

Corollary 2. *With $\lambda > {}^1\!/p_{\min}$, the algebraic soft-decoding radius exceeds the GS decoding radius if*

$$R < \left(\frac{p_{\min} - \frac{1}{\lambda}}{\sqrt{\gamma + \frac{1}{\lambda} - p_{\max} + p_{\min}}} \right)^2 . \tag{22}$$

Proof. Solving the equation

$$\frac{p_{\max} - \sqrt{R\left(\gamma + \frac{1}{\lambda}\right) - \frac{1}{\lambda}}}{p_{\max} - p_{\min}} > 1 - \sqrt{R}$$

for R and assuming $\lambda > {}^1\!/p_{\min}$ yields the corollary. □

¿From Corollary 2, we see that the ASD error radius exceeds the GS error radius for a subset of low rates for λ large enough. However, this subset may be small and not achievable if n is small. One can see this subset clearly in Example 1. One also notices in Example 1 that there is another non-zero subset of code rates where the transmitted codeword is always on the list, i.e. ${}^t\!/n \leq 1$. Corollary 3 quantifies this region.

Corollary 3. *Let $\lambda > {}^1\!/p_{\min}$. If*

$$R \leq \frac{\left(p_{\min} - \frac{1}{\lambda}\right)^2}{\gamma + \frac{1}{\lambda}}, \tag{23}$$

then an algebraic soft-decision decoder will always produce a list containing the transmitted codeword $\underline{c}$.

Proof. If the RHS of (13) is 1, then ASD will produce a list that contains $\underline{c}$ regardless of the error pattern. Thus, the task reduces to solving for R in the inequality

$$\frac{p_{\max} - \sqrt{R\left(\gamma + \frac{1}{\lambda}\right) - \frac{1}{\lambda}}}{p_{\max} - p_{\min}} \geq 1,$$

from which the corollary follows. □

It is a surprising result that there exists non-zero rates where ASD always produces a list that contains the transmitted codeword. The intuition in support of this result is that there is always a path from the transmitted vector to the received vector. Thus, the soft-decision decoder can neglect those codewords that could not have been transmitted due to zeros in the transition probability matrix, allowing the decoder to produce a list that is polynomial in n.

For the q-ary symmetric channel, a case where intuition tells us that ASD should provide no improvement over GS decoding, Figure 3.2 still shows a region that where ASD's error radius exceeds GS decoding's error radius. However, there is no code construction that simultaneously satisfies the conditions of Corollary 3 and (22) for the q-ary symmetric channel unless we have $\lambda \to \infty$. Thus, there is no achievable rate region where the ASD radius is larger than the GS decoding radius except when the size of the list is unbounded.

Even though the decoder always has the transmitted codeword on its list for rates that satisfy Corollary 3, there is still a non-zero probability of a decoding error due to another codeword also on the list being ultimately selected by the decoder. The probability of a wrong codeword being selected is a separate error event that will be discussed in more detail in Section 3.6.

3.5. **Multivariate Error Decoding Radius.** Suppose a PV code is transmitted over an discrete additive channel with error probabilities $\{p_1, p_2, ...p_{q^M}\}$. The statistics $p_{\min}$, $p_{\max}$, and γ are defined as before over this new set of transition probabilities. An error radius is given in Theorem 3 for soft-decision decoding of PV codes.

Theorem 3. *Given a PV code with rate $R = \frac{k}{Mn}$ is used to communicate over an additive-noise channel. If*

$$(24) \qquad \frac{t}{n} \leq \frac{p_{\max} - \sqrt[M+1]{\frac{R^M M^M}{(M+1)!} \sum_{i=1}^{q^M} \prod_{l=0}^{M} (p_i + l/\lambda)} - \frac{1}{\lambda}}{p_{\max} - p_{\min}},$$

then an algebraic soft-decision decoder, with complexity factor λ, will produce a list that contains the transmitted codeword $\underline{c}$.

The proof is done in exactly the same way as Theorem 2 and is omitted.

Example 4. *Let us return to the typewriter channel, $p_{\max} = 0.8$, $p_{\min} = 0.2$, $\gamma = 0.68$, and $\lambda = 100$, and compare trivariate soft-decision decoding of PV codes to bivariate soft-decision decoding of RS codes. Figure 3.4 shows the error radii (13) and (24). The graph shows that trivariate decoding provides an improvement over bivariate for rates less than 0.3.*

3.6. **Bound on the Probability of Decoding Error.** This section is focused on bounding the probability of error for Algebraic Soft-Decision Decoding. In the list-decoding setting, the probability of error has generally been defined as the probability that the transmitted codeword is not on the decoder's list. In many applications, in the final stage of decoding, it is required to select a unique codeword candidate from the list obtained using the maximum likelihood criteria. In this section, we derive a bound for the probability of error when the criteria is the transmitted codeword is on the decoder's list *and* it is selected as the best estimate.

The only known previous work in deriving ASD probability of error bounds is [15]. In that paper, Ratnakar and Koetter consider a general channel and take the probability of error to be the event that the transmitted codeword is not on the decoder's list. We propose a more comprehensive probability of error that includes the event that the decoder chooses the wrong codeword form the list, and we present an upper bound for this probability of error.

3.6.1. *General Form.* As in the previous section, the channel is assumed to be additive and memoryless. The transmitted codeword is $\underline{c}$, the decoder's list is $\mathcal{L}$, and the decoder's chosen codeword is $\hat{c}$. Define the following random events $\mathcal{A}$ and $\mathcal{B}$ as

$$\mathcal{A} : \underline{c} \notin \mathcal{L} \qquad \mathcal{B} : \{\hat{c} \neq \underline{c}\}$$

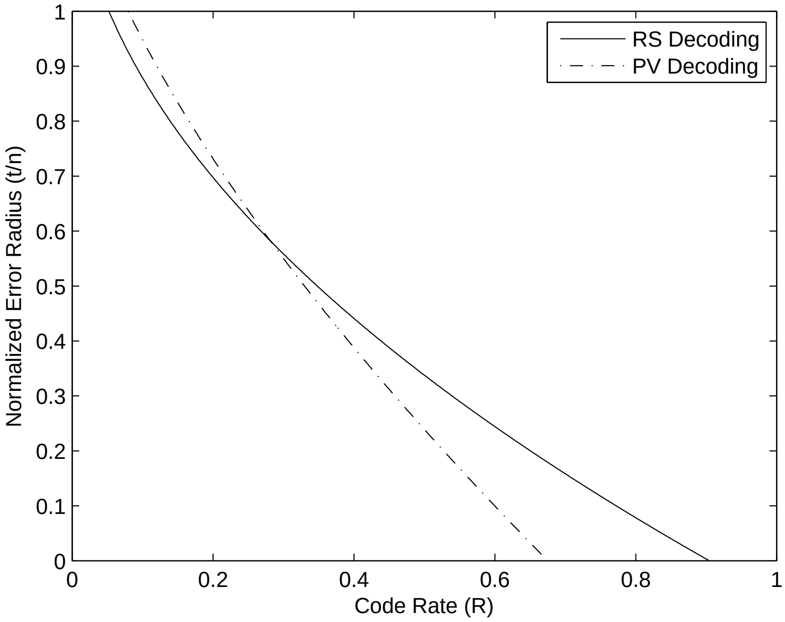

FIGURE 7. Trivariate decoding of a PV code compared to ASD of a RS code.

The list-decoding probability of error, which is the probability that the list produced by ASD contains the transmitted codeword, is given by $\Pr\{\mathcal{A}\}$. The selection probability of error is given by $\Pr\{\mathcal{B}\}$. Observe that $\mathcal{A} \subseteq \mathcal{B}$, giving us $\Pr\{\mathcal{A}\} \leq \Pr\{\mathcal{B}\}$. Each of these probabilities can be bounded above by using the Chernoff bound (see e.g. [5]).

Lemma 9. *(Chernoff Bound) Let w be a random variable with moment generating function $\Phi_w(s)$ and let A be a real number. Then*

$$\Pr\{w \geq A\} \leq e^{-sA}\Phi_w(s) \quad s > 0 \tag{25}$$

$$\Pr\{w \leq A\} \leq e^{sA}\Phi_w(-s) \quad s > 0. \tag{26}$$

In applications, one optimizes on the choice of s to obtain the tightest bound possible. The Chernoff Bound will allow us to write

$$\Pr\{\mathcal{A}\} \leq e^{-nE_{\mathcal{A}}}, \qquad \Pr\{\mathcal{B}\} \leq e^{-nE_{\mathcal{B}}}.$$

The functions $E_{\mathcal{A}}$ and $E_{\mathcal{B}}$ are the error exponents for $\Pr\{\mathcal{A}\}$ and $\Pr\{\mathcal{B}\}$, respectively.

3.6.2. *List-Decoding Probability of Error.* Theorem 4 gives an upper bound on the probability that the transmitted codeword is not on the decoder's list.

Theorem 4. *The probability of event $\mathcal{A}$ can be bounded above as*

$$\Pr\{\mathcal{A}\} \leq e^{-nE_{\mathcal{A}}},$$

where

$$E_{\mathcal{A}} = \infty \qquad if R < \frac{(p_{\min} - \frac{1}{\lambda})^2}{\gamma + \frac{1}{\lambda}}, \lambda > \frac{1}{p_{\min}}$$

$$E_{\mathcal{A}} = -\ln\left(\sum_{i=1}^{q} p_i e^{-s\left(p_i - \sqrt{R(\gamma+\frac{1}{\lambda})} - \frac{1}{\lambda}\right)}\right) \quad otherwise.$$

Proof. According to Corollary 2, we know $\underline{c} \in \mathcal{L}$ for all error patterns when

$$R \leq \frac{(p_{\min} - \frac{1}{\lambda})^2}{\gamma + \frac{1}{\lambda}}, \lambda > \frac{1}{p_{\min}}.$$

Thus, $E_{\mathcal{A}} = \infty$ in this rate region. In order to gain insight into the the event $\mathcal{A}$ for the remainder of the rates, consider again Lemma 8. If the condition in Lemma 8 is met, it is clear that $\mathcal{A}$ is false, but if the condition is not met, $\mathcal{A}$ could either be true or false. Thus, $\mathcal{A}$ will be true a certain fraction of the time when Lemma 8 is false. Now recall that (18) follows directly from Lemma 8, except it is generalized to all possible received vectors $\underline{y}$. Following the logic above, the probability that (18) is false will give an upper bound on $\Pr\{\mathcal{A}\}$. On account of Lemma 8, we obtain

$$\begin{aligned}\Pr\{\mathcal{A}\} &\leq \Pr\left\{\sum_{i=1}^{n} \mathbf{W}_i \leq n\sqrt{R\left(\gamma + \frac{1}{\lambda}\right) + \frac{n}{\lambda}}\right\} \\ &\leq e^{sn\left(\sqrt{R(\gamma+\frac{1}{\lambda})}+\frac{1}{\lambda}\right)} \left(\sum_{i=1}^{q} p_i e^{-sp_i}\right)^n \qquad (s > 0)\end{aligned}$$

which proves the theorem. □

In order to obtain the tightest bound, we need to maximize $E_{\mathcal{A}}$ through proper choice of s. If we define $g(s)$ by $E_{\mathcal{A}} = -\ln(g(s))$, then the goal is to minimize $g(s)$. When the maximum value of $E_{\mathcal{A}}$ is negative, then reliable communication is not possible. The following lemma shows that reliable communication is possible when the code rate is less than a rate maximum that can be well-estimated by γ.

Lemma 10.

$$\begin{aligned}&E_{\mathcal{A}} > 0 \qquad \textit{if} \quad R < R_{\max} = \frac{(\gamma - \frac{1}{\lambda})^2}{\gamma + \frac{1}{\lambda}} \\ &E_{\mathcal{A}} = 0 \qquad \textit{otherwise.}\end{aligned}$$

Proof. We have that

$$g(s) = \sum_{i=1}^{q} p_i e^{-s\left(p_i - \sqrt{R(\gamma+\frac{1}{\lambda})}-\frac{1}{\lambda}\right)}.$$

First observe that $g''(s) > 0$, indicating that the function $g(s)$ is convex. Next observe that $g(0) = 1$ and that $g'(0) < 0 \leftrightarrow R < \frac{(\gamma-\frac{1}{\lambda})^2}{\gamma+\frac{1}{\lambda}}$. In the case that $g'(0) < 0$, the minimum value of $g(s)$ is achieved for some $s' > 0$, and since $g(0) = 1$, $g(s') \leq 1$. Thus, $E_{\mathcal{A}} > 0$. Otherwise, $g(s) \geq 1$ which implies $E_{\mathcal{A}} = 0$. □

3.6.3. *Probability of Error for Selection.* When the list-decoding probability of error is zero, one does not have insight into the performance of ASD. Since the transmitted codeword is always on the decoder's list, the probability of decoder error only exists in the selection phase. Therefore, it is of interest to quantify a comprehensive probability of error given by $\Pr\{\mathcal{B}\}$ and bounded in Theorem 5.

Theorem 5. *Let C be a q-ary code of length n and rate R. The probability of event $\mathcal{B}$ can be bounded above as*

$$\Pr\{\mathcal{B}\} \leq e^{-nE_{\mathcal{B}}}$$

for at least one coset of a RS code, where

$$E_{\mathcal{B}} = -\ln\left(q^{R-1}e^{s/\lambda} + 2q^{R-1}\sum_{i=1}^{q}\sum_{\substack{j=1\\j\neq i}}^{q} p_i e^{-s(p_i-p_j-\frac{1}{\lambda})}\right).$$

Proof. Consider the following random code construction. Select a RS code C of rate R, and denote the cosets of the code as C_i, $i = 1, \ldots\ q^{n-k}$ with any ordering. Suppose a message is transmitted using the random code C' defined as

$$\Pr\{C' = C_i\} = \frac{1}{q^{n-k}}. \tag{27}$$

If the probability of error can be bounded above for this random-code construction, then there is at least one coset of C that has a probability of error with this same upper bound. Observe that

$$\Pr\{\mathcal{B}\} = \Pr\{\exists \underline{c}' \neq \underline{c},\ \underline{c}' \in \mathcal{L} : S_M(\underline{c}') \geq S_M(\underline{c})\} \tag{28}$$

Define the events $\mathcal{C}_i$ and $\mathcal{D}_i$ as follows:

$$\mathcal{C}_i : \{(\underline{c}_i \in \mathcal{L}) \quad \& \quad (S_M(\underline{c}_i) \geq S_M(\underline{c}))\}$$
$$\mathcal{D}_i : \{S_M(\underline{c}_i) \geq S_M(\underline{c})\}$$

Clearly, $\Pr\{\mathcal{D}_i\} \geq \Pr\{\mathcal{C}_i\}$. Assume that the codewords of each coset are ordered in some particular way of which $\underline{c}_i$ is a member codeword. Next, continuing from (28)

$$\begin{aligned}\Pr\{\exists \underline{c}' \neq \underline{c},\ \underline{c}' \in \mathcal{L} : S_M(\underline{c}') \geq S_M(\underline{c})\} &= \Pr\Big\{ \bigcup_{\substack{i=1\\ i:\underline{c}_i \neq \underline{c}}}^{q^k} \mathcal{C}_i \Big\} \leq \sum_{\substack{i=1\\ i:\underline{c}_i \neq \underline{c}}}^{q^k} \Pr\{\mathcal{C}_i\} \\ &\leq \sum_{\substack{i=1\\ i:\underline{c}_i \neq \underline{c}}}^{q^k} \Pr\{\mathcal{D}_i\}. \end{aligned} \tag{29}$$

Define the event $\mathcal{E}_i$ as follows:

$$\mathcal{E}_i : \{\underline{v}_i \in \mathbb{F}_q^n : S_M(\underline{v}_i) \geq S_M(\underline{c})\}.$$

The probability that a given vector is a codeword in C_i is $\frac{1}{q^{n-k}}$. Using this fact, (29) can be rewritten

$$\begin{aligned}\sum_{\substack{i=1\\ i:\underline{c}_i \neq \underline{c}}}^{q^k} \Pr\{\mathcal{D}_i\} &= \sum_{\substack{i=1\\ i:\underline{v}_i \neq \underline{c}}}^{q^n} \Pr\{\mathcal{E}_i, \underline{v}_i \in C'\} \\ &= \sum_{\substack{i=1\\ i:\underline{v}_i \neq \underline{c}}}^{q^n} \frac{1}{q^{n-k}} \Pr\{\mathcal{E}_i\}. \end{aligned} \tag{30}$$

For a given message, the codeword transmitted is drawn uniformly from a set of candidate codewords that are equally likely to have any of the q symbols at each symbol position. Thus, one element in $S_M(\underline{v}_i)$ is drawn uniformly from the set of q possible multiplicities. Let $\mathbf{V}_i$ be a random variable uniformly distributed on $\{p_1, p_2, \ldots p_q\}$. We have

$$\begin{aligned}\sum_{\substack{i=1\\ i:\underline{v}_i \neq \underline{c}}}^{q^n} \frac{1}{q^{n-k}} \Pr\{\mathcal{E}_i\} &= \frac{q^n - 1}{q^{n-k}} \Pr\Big\{ \sum_{i=1}^{n} \lfloor \lambda \mathbf{V}_i \rfloor \geq \sum_{i=1}^{n} \lfloor \lambda \mathbf{W}_i \rfloor \Big\} \\ &\leq q^k \Pr\Big\{ \sum_{i=1}^{n} \big(\lambda \mathbf{V}_i - \mu_{i,V}\big) \geq \sum_{i=1}^{n} \big(\lambda \mathbf{W}_i - \mu_{i,W}\big) \Big\} \\ &\leq q^k \Pr\Big\{ \sum_{i=1}^{n} \big(\mathbf{V}_i - \mathbf{W}_i\big) \geq -\frac{n}{\lambda} \Big\}. \end{aligned}$$

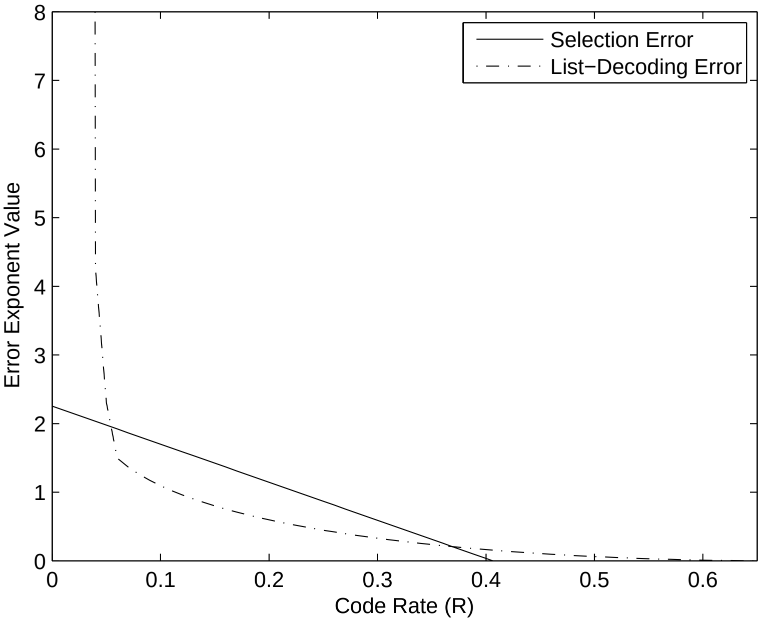

FIGURE 8. Comparison of the two error exponents for ASD over low rates.

The second line introduces PMAS error terms, that satisfy $0 \leq \mu_{i,V}, \mu_{i,W} < 1$, similarly to the proof of Theorem 1. The V and W subscripts indicate the random variable to which the error terms are associated. We arrive at the third line by setting the error terms to their extreme points that make the probability expression greater. Now let us apply the Chernoff bound (25). For any $s > 0$,

$$\Pr\Big\{\sum_{i=1}^{n}(\mathbf{V}_i - \mathbf{W}_i) \geq -\frac{n}{\lambda}\Big\} \leq e^{\frac{sn}{\lambda}}\Big(\sum_{i=1}^{q}\frac{1}{q}e^{sp_i}\Big)^n\Big(\sum_{i=1}^{q}p_i e^{-sp_i}\Big)^n. \tag{31}$$

Algebraic manipulation of (31) and the appropriate definition of $E_{\mathcal{B}}$ results in the theorem. □

Define $E_{\mathcal{B}} = -\ln(h(s))$. Minimizing the bound is the same as minimizing $h(s)$, and this procedure is independent of the value of R. It is of interest to find out the behavior of $E_{\mathcal{A}}$ compared to $E_{\mathcal{B}}$ across all values of R. For the same channel in Example 1 ($p_{\max} = 0.8$, $p_{\min} = 0.2$, $\lambda = 100$, $q = 256$, and $\gamma = 0.68$), Figure 3.5 shows the comparison of $E_{\mathcal{A}}$ to $E_{\mathcal{B}}$. The One can see that reliable communication, in the list-decoding sense, is assured for rates less than 0.67, and reliable communication, based on the probability of error of selection, is guaranteed for rates less than 0.41.

4. CONCLUSION

The results presented in this thesis have shown that Algebraic Soft-Decision Decoding is able to outperform its hard-decision counterparts for low-rate to medium-rate codes. The error decoding radius presented is a new result in the literature that allows ASD to be compared to other RS decoding methods. A comprehensive probability of error bound is derived that includes the previously overlooked probability of selection error.

4.1. **Open Questions.** Though there has generally been a dearth of publications related to Algebraic Soft-Decision Decoding, one area of relatively high activity has been the investigating the best method of generating the multiplicity matrix from the channel. The PMAS method was chosen to use in this thesis, but it is unknown if other methods may yield a larger decoding radius.

Throughout this thesis, it has been assumed that the channel is discrete. An interesting question is if an error radius can be derived for the Gaussian channel. Techniques used in this thesis do not seem to be sufficient to accomplish this goal. Another possible avenue of research is a study of the trade-off among the error-correction radius, error probability, and the size of the decoder's list. It also would be worthwhile to extend ASD to the decoding of Algebraic-Geometry codes.

An open question that remains unanswered is if ASD's performance makes it a worthwhile decoder to use in Reed-Solomon coding applications. For low-rate coding applications with channels that are far from q-ary symmetric, ASD shows the potential to correct a greater number of errors than hard-decision decoders. However, there is no insight in this thesis about the performance of high-rate ASD decoding of high-rate codes.

We are interested in decoding high-rate codes because they are the most frequently used to solve real-world engineering problems. Koetter and Vardy have projected in [11], through simulation, that a 1 dB coding gain can be achieved over GS and BM decoding for a RS code of rate 0.92. If the channel used in the simulation was q-ary symmetric, then this coding gain would likely be reduced, but the question remains of what is the improvement that ASD gives the user in high-rate coding applications.

REFERENCES

[1] R. E. Blahut, Theory and Practice of Error-Correcting Codes, Addision-Wesley, 1983.

[2] T. Cover and J. Thomas, Elements of Information Theory, John Wiley and Sons, 2001.

[3] M. El-Khamy and R. McEliece, “Interpolation Multiplicity Assignment Algorithms for Algebraic Soft-Decision Decoding of Reed-Solomon Codes,” DIMACS Series in Algebraic Coding Theory and Information Theory, A. Ashikhmin and A. Barg (Eds.), pp. 99-120, AMS, 2005.

[4] G. L. Feng and K. K. Tzeng, “A Generalization of the Berlekamp-Massey Algorithm for Multisequence Shift-Register Synthesis with Applications to Decoding Cyclic Codes,” IEEE Trans. Inform. Theory, vol. **IT-37**, no. 5, pp. 1274-1287, Sep. 1991.

[5] R. Gallager, Information Theory and Reliable Communication, John Wiley and Sons, 1968.

[6] W. Gross, F. Kschischang, R. Koetter, and G. Gulak, “Applications of Algebraic Soft-Decision Decoding of Reed-Solomon Codes,” preprint.

[7] V. Guruswami and M. Sudan, “Improved Decoding of Reed-Solomon and Algebraic-Geometric Codes,” IEEE Trans. Inform. Theory, vol. **IT-45**, pp. 1757-1767, Sept. 1999.

[8] J. Jiang and K. Narayanan, “Performance Analysis of Algebraic Soft Decoding of Reed-Solomon Codes over Binary Symmetric and Erasure Channels,” Proceedings of the 2003 IEEE International Symposium on Information Theory, p. 1186-1190.

[9] J. Justesen, “Soft-Decision Deocding of RS Codes,” Proceedings of the 2005 IEEE International Symposium on Information Theory, p. 1183-1185.

[10] R. Koetter, “Fast Generalized Minimum-Distance Decoding of Algebraic-Geometry and Reed-Solomon Codes,” IEEE Trans. Inform. Theory, vol. **IT-42**, no. 3, pp. 721-736, May 1996.

[11] R. Koetter and A. Vardy, “Algebraic Soft-Decision Decoding of Reed-Solomon Codes,” IEEE Trans. Inform. Theory, vol. **IT-49**, pp. 2809-2825, Nov. 2003.

[12] R. McEliece, “The Guruswami-Sudan Decoding Algorithm for Reed-Solomon Codes,” 2003, manuscript, available on-line at http://www.systems.caltech.edu/EE/Faculty/rjm/.

[13] F. Parvaresh and A. Vardy, “Multiplicity Assignments for Algebraic Soft-Decision of Reed-Solomon Codes,” Proceedings of the 2003 IEEE International Symposium on Information Theory, p. 250.

[14] F. Parvaresh and A. Vardy, “Correcting Errors Beyond the Guruswami-Sudan Radius in Polynomial Time,” Proceedings of the 2005 IEEE Annual Symposium on the Foundations of Computer Science (FOCS), pp. 246-257, 2005.

[15] N. Ratnakar and R. Koetter, “Exponential Error Bounds for Algebraic Soft-Decision Decoding of Reed-Solomon Codes,” IEEE Trans. Inform. Theory, vol. **IT-51**, pp. 3899-3917, Nov. 2005.

[16] R. Roth and G. Ruckenstein, “Efficient Decoding of Reed-Solomon Codes Beyond Half the Minimum Distance,” IEEE Trans. Inform. Theory, vol. **IT-46**, no. 1, pp. 246-257, Jan. 2000.

[17] C. Shannon, “A Mathematical Theory of Communication ,” Bell System Technical Journal, vol. **27**, pp. 379-423 (Part I) and 623-656 (Part II), 1948.