

# Performance Analysis of Algebraic Soft-Decision Decoding of Reed-Solomon Codes

Andrew Duggan\* and Alexander Barg§

**Abstract**—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 correction 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 (GS) hard-decision decoding for a subset of low-rate codes. These results are also extended to multivariable interpolation in the sense of Parvaresh and Vardy. An upper bound is then presented 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 than the results available in the literature.

## I. INTRODUCTION

Reed-Solomon (RS) codes are used in a wide variety of applications currently, and the classical algorithm of Berlekamp and Massey (BM algorithm) has been employed for decoding in most cases. For an RS code of length  $n$  and dimension  $k$ , this algorithm is guaranteed to recover the transmitted codeword within an error radius of  $\lfloor \frac{n-k+1}{2} \rfloor$ .

Guruswami and Sudan [7] presented an important new algebraic decoding method for RS codes that is able to correct errors beyond the BM decoding radius. This method involves constructing a bivariate polynomial with zeros of multiplicity based on the received symbols. The polynomial can then be factored to give a list of candidate codewords; thus, it is a list decoder. A Guruswami-Sudan (GS) decoder includes the transmitted codeword on its output list if the errors fall within a radius of  $n - \sqrt{nk}$ .

In [7], the authors mention that their hard-decision algebraic decoding technique can be extended to soft-decision decoding by setting the values of the multiplicities based on channel posterior probabilities and not received symbols. However, [7] does not provide a way of assigning these multiplicities, which turns out to be a nontrivial component in the RS decoding procedures. Koetter and Vardy refined the Algebraic Soft-Decision Decoding (ASD) approach in [11] by providing an algorithm that converts a matrix of posterior probabilities  $\Pi$  to a multiplicity assignment matrix  $\mathcal{M}$ .

Various papers, such as [13], [3], [8], have been published since [11] that propose using a different method for converting  $\Pi$  to  $\mathcal{M}$ . However, few papers have given insight into the decoding region of ASD. One exception is [8], which characterizes the decoding region for medium to high rate codes over

binary erasure and binary symmetric channels. Another paper [10] derives an error correction radius for an arbitrary additive cost function associated with transitions in the channel.

This paper examines the performance of ASD when the noise is additive, i.e., there is a probability distribution  $p(a)$  on  $q$ -ary errors that does not depend on the transmitted sequence. Relying on this distribution, we derive in Sect. III simple estimates of the error radius within which the transmitted codeword is guaranteed to be on the list produced by ASD. With the simple expression for the error radius obtained, we are also able to characterize the region where the ASD algorithm provides an improvement over the GS decoding radius.

In Section IV, we study bounds on the probability of error for ASD decoding. Prior work [15] has concentrated on the list-decoding error event, namely the event that the transmitted codeword fails to be included in the list. We point out that for low code rates, the list decoding error criterion does not provide insight into the performance of the decoder because the transmitted codeword always will be included in the list. We therefore define decoding to be successful if the ASD algorithm selects the transmitted codeword from the decoder's list, and we derive an upper bound for the probability of error based on this new definition. Finally, in Section V we briefly discuss soft-decision decoding of multivariate RS codes introduced in a recent work of Parvaresh and Vardy [14] and extend to this case our estimate of the ASD error radius.

The error bounds obtained in Sections III-V are either the first of their kind available in the literature, or, as in the case of the ASD list-decoding error bound, improve the results known previously.

## II. DECODING OF RS CODES

### A. Notation

Let  $q$  be a prime power, let  $\mathbb{F}_q = \{\alpha_1 = 0, \alpha_2, \dots, \alpha_q\}$  be the finite field of  $q$  elements, and let  $n = q - 1$  be the code length. Denote  $\text{dist}(\cdot, \cdot)$  as the Hamming distance. With a vector  $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{F}_q^n$  we associate an integer-valued function  $v(i)$  whose value is  $j$  if  $v_i = \alpha_j$ .

For a polynomial  $f \in \mathbb{F}_q[x]$  define the evaluation mapping  $\text{eval} : f \rightarrow \mathbb{F}_q^n$  given by  $(\text{eval} f)_i = f(\alpha_i)$ ,  $2 \leq i \leq q$ . 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

$$\{c = \text{eval}(f) : f \in \mathbb{F}_q[x], 0 \leq \deg f \leq k - 1\}.$$

To describe the encoding of the code  $C$ , suppose that the message to be transmitted is  $\mathbf{u} = (u_1, u_2, \dots, u_k)$  where  $u_i \in$

Presented in part at the 44th Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, September 27-29, 2006.

\* E-mail address: aduggan@alum.mit.edu.

§ Dept. of ECE and Institute for Systems Research, University of Maryland, College Park, MD 20742, abarg@umd.edu. Supported in part by NSF grants CCF0515124, CCF0635271, and by NSA grant H98230-06-1-0044.

$\mathbb{F}_q$ ,  $1 \leq i \leq k$ . The codeword that corresponds to it is given by  $\mathbf{c} = \text{eval}(f)$ , where the polynomial  $f$  has the form

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

We assume that the codeword  $\mathbf{c}$  is transmitted over a discrete memoryless channel. In the hard-decision case, the output of the channel is the vector  $\mathbf{y} = \mathbf{c} + \mathbf{e}$ . Let  $w_{i,j} = \Pr(y = \alpha_i | c = \alpha_j)$  be the probability that the symbol  $\alpha_j$  transmitted over the channel is received as  $\alpha_i$ . We will assume that the noise is *additive*, i.e. there exists a probability distribution  $p$  on  $\mathbb{F}_q$  such that  $\Pr(\alpha + e | \alpha) = p(e)$ ,  $\alpha, e \in \mathbb{F}_q$ . Note that under these assumptions, the channel is symmetric as defined in Section 8.2 of [2].

In the setting of soft-decision decoding, the demodulator is assumed to provide the decoder with the posterior probabilities conditioned on the received (continuous) signal. However, the task of analyzing this setting in the context of algebraic list decoding so far has proved elusive. We will therefore assume that the receiver outputs in each position of the codeword, a hard-decision symbol (for instance, the most likely transmitted symbol) *together* with the  $q$  values of *posterior probabilities*  $\pi_{i,j} = \Pr(c = \alpha_i | y = \alpha_j)$ . As customary in the literature, we will assume that ASD takes the channel's output to be in the form of a  $q \times n$  matrix  $\Pi = [\pi_{i,j}]$ .

This paper concentrates on the ASD algorithm of [11] and compares its performance to the well-known hard-decision decoding algorithms of Berlekamp and Massey (see, e.g., [1]) and Guruswami and Sudan [7], [12].

### B. Hard-Decision Decoding Methods

Under BM decoding, if the number of errors  $t = \text{dist}(\mathbf{c}, \mathbf{y})$  satisfies

$$t \leq \left\lfloor \frac{n - k + 1}{2} \right\rfloor, \quad (1)$$

then the decoder will output  $\mathbf{c}$ . If condition (1) is not true, then decoding is guaranteed to fail. Therefore, (1) is a necessary condition for BM decoding success.

GS decoding produces a list that contains all the codewords of a certain distance  $t_m$  from the vector  $\mathbf{y}$  and potentially some codewords outside of this Hamming ball. List decoding success is declared if the correct codeword is on the list. 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  $m \rightarrow \infty$ . Let  $\mathbf{c}$  be a codeword that satisfies

$$\text{dist}(\mathbf{y}, \mathbf{c}) < n - \sqrt{nk}. \quad (2)$$

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

The complexity of the algorithm often becomes a limiting factor before the maximum possible  $t_m$  is achieved. Note that (2) is only a sufficient condition on GS list-decoding success: the decoding is guaranteed to have the transmitted codeword on the list if the error pattern satisfies (2), assuming large  $m$ .

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

$$\begin{aligned} \tau &= \frac{1 - R}{2} \quad (\text{BM Decoding}) \\ \tau &= 1 - \sqrt{R} \quad (\text{GS Decoding, } m \text{ large}). \end{aligned} \quad (3)$$

The two error radii given in (3) are shown as the dashed curves in Figure 1(a). The GS decoding radius is always greater than its BM counterpart although the difference becomes small for high rates. However, no error radius for Algebraic Soft-Decision Decoding was previously known.

### C. ASD Algorithms

ASD is an extension of GS decoding by the manipulation of the multiplicities. We only mention the salient features of this algorithm, referring to [11] for the details. Instead of operating on a vector  $\mathbf{y}$ , ASD takes as input a multiplicity matrix  $\mathcal{M}$  of dimensions  $q \times n$  constructed on the basis of the posterior probability matrix  $\Pi$ . The soft-decision decoder constructs a bivariate polynomial  $Q(X, Y)$  that has zeros of multiplicity set by  $\mathcal{M}$ . In the next step, the decoder recovers a list  $\mathcal{L}$  of putative transmitted codewords from the  $Y$ -zeros of  $Q(X, Y)$ . In contrast to GS decoding that always constructs  $Q(X, Y)$  based on  $n$  distinct zeros, ASD can have up to  $qn$  distinct zeros.

*Definition 2:* Define the *Score* of a vector  $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{F}_q^n$  with respect to a multiplicity matrix  $\mathcal{M}$  to be

$$S_{\mathcal{M}}(\mathbf{v}) = \sum_{i=1}^n m_{v(i), i}.$$

where  $v(i) = \ell$  if  $v_i = \alpha_\ell$ .

If needed, in the last stage the decoder chooses the codeword  $\mathbf{c}$  from the list  $\mathcal{L}$  with the largest score or the codeword with the largest probability  $P^n(\mathbf{c} | \mathbf{y}) = \prod \pi_{c_i, y_i}$ .

1) *The Multiplicity Matrix:* The matrix  $\mathcal{M}$  is determined from the matrix of posterior probabilities  $\Pi$ . Koetter and Vardy [11], Parvaresh and Vardy [13], and El-Khamy and McEliece [3] have proposed various methods for determining  $\mathcal{M}$  from  $\Pi$ . A simple method for converting  $\Pi$  to  $\mathcal{M}$  that will be used in this paper is called the Proportionality Multiplicity Assignment Strategy (PMAS) proposed by Gross et al. [5]. PMAS finds  $\mathcal{M}$  by performing the following element-wise calculation on  $\Pi$  for some fixed number  $\lambda$ :

$$m_{i,j} = \lfloor \lambda \pi_{i,j} \rfloor, \quad i = 1, \dots, 1; j = 1, \dots, n. \quad (4)$$

Thus, the matrix  $\mathcal{M}$  is determined uniquely from the received vector  $\mathbf{y}$  and the properties of the communication channel. The parameter  $\lambda \in \mathbb{Z}^+$  is the complexity factor, and its adjustment controls directly 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, defined as follows.

*Definition 3:* Let the *Cost* of a multiplicity matrix  $\mathcal{M}$  be

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

2) *Threshold Condition*: GS decoding includes on its output list all codewords in the Hamming space within a certain radius of the received vector  $\mathbf{y}$ , but ASD has no known geometric interpretation. A sufficient condition for ASD's success is determined indirectly by the vector  $\mathbf{y}$  and can be stated in terms of the score  $S_{\mathcal{M}}(\mathbf{c})$  as given in the next lemma.

*Lemma 2*: [11] Suppose ASD is used to decode a received vector  $\mathbf{y}$  with the RS code. If

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

or equivalently

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

then the transmitted codeword is on the decoder's list.

Lemma 2 can be used to evaluate ASD's performance in the case that  $\mathbf{c}$  is transmitted,  $\mathbf{y}$  is received, and  $\mathcal{M}$  is the multiplicity matrix.

3) *Size of the List*: The decoder's list  $\mathcal{L}$  cannot exceed the  $Y$ -degree of  $\mathcal{Q}(X, Y)$  since  $\mathcal{L}$  is obtained by factoring  $\mathcal{Q}(X, Y)$  for the  $Y$ -roots. The size of the list in terms of the  $Y$ -degree  $L$  of  $\mathcal{Q}$  is estimated in the next lemma, which is due to McEliece [12], Corollary 5.14.

*Lemma 3*: For  $k \geq 2$ , the number of codewords on the list of an algebraic soft-decision decoder does not exceed

$$L = \left\lfloor \sqrt{\frac{2\mathcal{C}(\mathcal{M})}{k-1} + \left(\frac{k+1}{2k-2}\right)^2} - \left(\frac{k+1}{2k-2}\right) \right\rfloor.$$

The condition  $|\mathcal{L}| \leq L$  is met with equality if and only if all the  $Y$ -roots have degree less than  $k$ .

### III. ASD ERROR CORRECTION PERFORMANCE

In this section, we present one of our main results, an estimate of the error correction radius  $t$  of the algorithm. We would like to stress one essential difference of the result below from the other similar results in the literature. In the case of BM and GS decoding for instance, all of the codewords within the error radius are included in the list output by the decoder. In contrast, we only guarantee in the case of ASD that if the *transmitted* codeword is distance  $t$  away from the received one, then it will be included on decoder's list. Other codewords, even within the sphere of radius  $t$  from  $\mathbf{y}$ , may escape being output by the decoder. In other words, the decoding regions of ASD are far from being spherical, and in fact, no geometric characterization of them is available.

#### A. Setting

Following Koetter and Vardy in [11] and Justesen in [9], we assume that each symbol entering the channel is uniformly drawn from  $\mathbb{F}_q$ . It follows that  $\pi_{i,j} = w_{i,j}$ . Since the channel is symmetric, we know that the channel transition probabilities  $w_{i,j}$  are drawn from the set  $\{p_1, p_2, \dots, p_q\}$ . Next, we will introduce the three channel statistics

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

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  $\gamma$  will be the only channel statistics necessary in our analysis of ASD's performance.

#### B. Error Radius

*Theorem 1*: Suppose that an RS code with rate  $R = k/n$  is used to communicate over an discrete, additive-noise channel. Suppose that a codeword  $\mathbf{c}$  is transmitted and an algebraic soft-decision decoder with complexity factor  $\lambda$  is used to decode the received vector  $\mathbf{y}$ . Let  $t = \text{dist}(\mathbf{c}, \mathbf{y})$ . If

$$\frac{t}{n} \leq \frac{p_{\max} - \sqrt{R(\gamma + \frac{1}{\lambda})} - \frac{1}{\lambda}}{p_{\max} - p_{\min}}, \quad (5)$$

then  $\mathbf{c}$  will be contained in the output list of the decoder.

*Proof*: Let  $\mathbf{c}$  be the transmitted codeword and let  $\mathbf{y}$  be the received vector. Substituting (4) in Lemma 2, we get

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

Instead of this condition for successful decoding let us use a more stringent one obtained by removing the integer parts:

$$\frac{\sum_{i=1}^n (\lambda w_{c(i),y(i)} - 1)}{\sqrt{\sum_{i=1}^n \sum_{j=1}^q ((\lambda w_{i,j})^2 + \lambda w_{i,j})}} \geq \sqrt{k-1}. \quad (6)$$

Rearranging and using the channel statistics  $\gamma$  yields

$$\frac{1}{n} \sum_{i=1}^n w_{c(i),y(i)} \geq \sqrt{R \left( \gamma + \frac{1}{\lambda} \right) - \frac{\gamma}{n} - \frac{1}{\lambda n} + \frac{1}{\lambda}}. \quad (7)$$

Inequality (7) certainly holds true for the codeword  $\mathbf{c}$  if

$$\frac{1}{n} \sum_{i=1}^n w_{c(i),y(i)} \geq \sqrt{R \left( \gamma + \frac{1}{\lambda} \right) + \frac{1}{\lambda}}. \quad (8)$$

We have derived a condition based on a specific  $\mathbf{y}$ , but we are interested in ASD's performance for any  $\mathbf{y}$  when  $\mathbf{c}$  is transmitted. Thus,  $\mathbf{y}$  becomes a random variable. Let  $W_i$  be a random transition probability  $w_{c(i),y(i)}$  given random  $\mathbf{y}$ . Note that the  $W_i$ s do not depend on  $\mathbf{c}$  because of the additive noise assumption. The p.m.f. of  $W_i$  is as follows:

$$p_{W_i}(p_i) := \Pr\{W_i = p_i\} = p_i, \quad i = 1, 2, \dots, q.$$

We can now rewrite (8) as follows:

$$\frac{1}{n} \sum_{i=1}^n W_i \geq \sqrt{R \left( \gamma + \frac{1}{\lambda} \right) + \frac{1}{\lambda}}.$$

Since the received vector  $\mathbf{y}$  differs from  $\mathbf{c}$  in  $t$  coordinates, we can bound the left-hand side of the last inequality below as follows:

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

Indeed, if  $y_i \neq c_i$  then the probability  $w_{c(i),y(i)} \geq p_{\min}$ . On the other hand, if there is no error in coordinate  $i$  then

the probability  $w_{c(i),y(i)} = p_{\max}$ ; otherwise, communication over the channel would be impossible. Thus,  $c$  is on the soft-decision decoder's list if

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

The theorem follows.  $\square$

A first look at the error radius in Theorem 1 reveals that (5) 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  $\lambda$  is sufficiently large, then the bound in Theorem 1 reduces to  $1 - \sqrt{R}$  which is the GS normalized error bound. Let us consider a few 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 1(a) shows the normalized error correction radius 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  $c$  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 below in this section.

*Example 2:* Next, let us look at the error radius of Theorem 1 when there are two possible, equiprobable errors per symbol transmission, termed the “two-error channel.” In this case,  $p_{\max} = 0.8$ ,  $p_{\min} = 0.1$ ,  $\gamma = 0.66$ , and we again set  $\lambda = 100$ . Figure 1(b) shows the ASD, GS, and BM curves for this example.

*Example 3:* Figure 1(c) 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  $\lambda = 100$ . 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.

As the channel approaches  $q$ -ary symmetric, ASD is shown to have a progressively smaller improvement over GS decoding for low-rate codes. The fact that the bound on the error radius for ASD is below the GS radius for higher rates is due to the approximations used in the proof of Theorem 1. The true ASD error radius is likely to be slightly greater or within a neighborhood of the GS decoding radius for all code rates.

In [10], Koetter considers a general transmission scenario when an RS or related code is used for communication over a symmetric channel, and the decoder's performance is measured in terms of an arbitrary additive cost function  $\Delta : \mathbb{F}_q \rightarrow \mathbb{R}^+$ . [10] attempts at optimizing the assignment of multiplicities for the purpose of finding the maximum attainable decoding radius. An error vector  $e \in \mathbb{F}_q^n$  is assigned the cost

$$\Delta(e) = \sum_{i=1}^n \Delta(e_i).$$

The result of [10] stated for an additive channel, is as follows (the original text contains some misprints).

*Theorem 2:* Suppose that a code vector  $c$  of an RS code of rate  $R$  transmitted over the channel is received as  $y = c + e$ . If the values  $R$  and  $\Delta(e)$  simultaneously satisfy the inequalities

$$R \leq (1 - \epsilon) \sum_{a \in \mathbb{F}_q} [\rho - \theta \Delta(a)]_+^2 \quad (\epsilon > 0) \quad (9)$$

$$\Delta(e) \leq \sum_{a \in \mathbb{F}_q} \Delta(a) [\rho - \theta \Delta(a)]_+ \quad (10)$$

where  $[x]_+ = \max(0, x)$ ,  $\theta > 0$  is an arbitrary parameter, and

$$\sum_{a \in \mathbb{F}_q} [\rho - \theta \Delta(a)]_+ = 1,$$

then the vector  $c$  will be placed on the ASD output list. Moreover, relations (9)-(10) characterize the optimum tradeoff between the code rate and the cost of error.

We note that even though the theorem's result is optimum for the “error radius” in terms of the cost function, its result is inferior to the result of Theorem 1 in the following sense. For the previous theorem to give a performance curve over Hamming distance, the decoder is optimized over Hamming distance, a concept that is congruent with a hard-decision decoder, not a soft-decision decoder. Namely, taking the distance as the cost function, we have  $\Delta(\alpha_i) = 1 - \delta_{\alpha_i, 0}$ ,  $i = 1, \dots, q$  (the cost is 1 for all the symbols of  $\mathbb{F}_q$  except for some distinguished symbol whose value is of no importance). Then (9)-(10) reduce to the equation

$$\begin{aligned} \frac{t}{n} &\leq \frac{n}{n+1} - \sqrt{\frac{nR}{(n+1)(1-\epsilon)} - \frac{n}{(n+1)^2}}, \\ \frac{1-\epsilon}{n+1} &\leq R \leq 1-\epsilon \end{aligned}$$

which is an error radius that slightly exceeds the GS bound (2) for most rates. As expected, it is inferior to (5) for a subset of low to medium rate codes.

### C. Size of the List

*Proposition 1:* Assume that  $k \geq 2$ . Given an additive noise channel, the size of the list for an algebraic soft-decision decoder is bounded above by

$$\left\lceil \sqrt{\frac{n(\lambda^2\gamma + \lambda)}{k-1} + \left(\frac{k+1}{2k-2}\right)^2} - \left(\frac{k+1}{k-2}\right) \right\rceil. \quad (11)$$

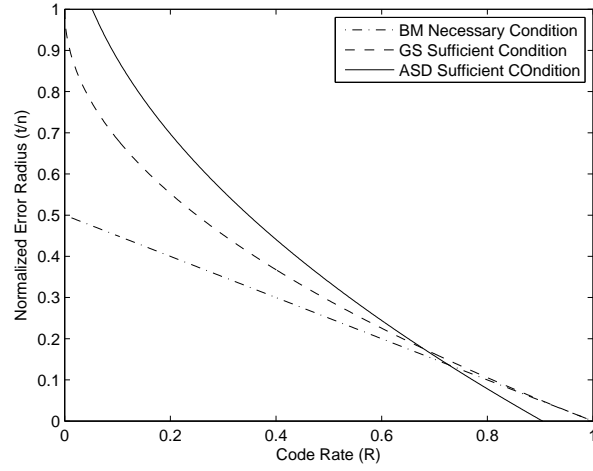
*Proof:* Assuming PMAS (4), observe that

$$2\mathcal{C}(\mathcal{M}) = n \sum_{i=1}^q [\lambda p_i]^2 + n \sum_{i=1}^q [\lambda p_i] \leq n\lambda^2\gamma + n\lambda.$$

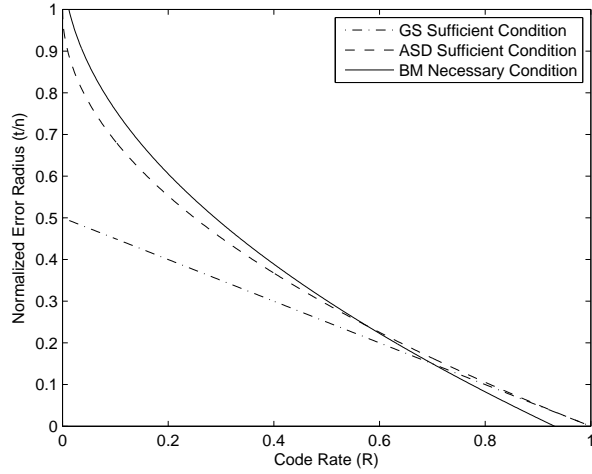
Substituting this upper bound in Lemma 3 gives the result.  $\square$

The bound in Proposition 1 is based on the  $Y$ -degree of the polynomial  $Q(X, Y)$ . Thus, it includes polynomials that are a good fit for the set of multiplicity points, including higher-degree polynomials. As a result, the bound of Proposition 1 is not tight.

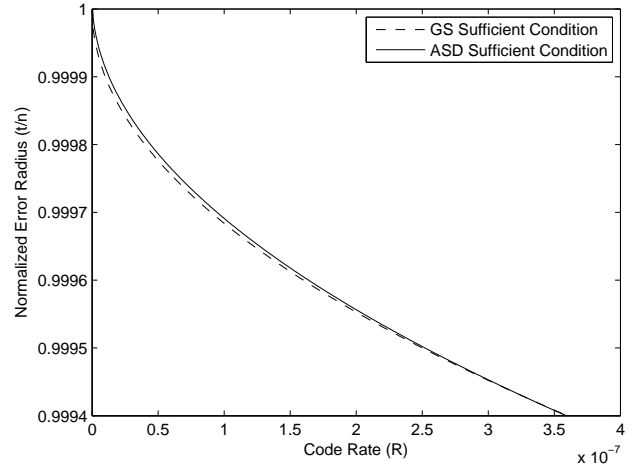
*Example 4:* Figure 2 shows a graph of the bound on the list size presented in Proposition 1 for Example 1 with  $n = 255$ .



(a)



(b)



(c)

Fig. 1. Decoding radius of ASD compared to GS and BM decoding for a), the “typewriter channel” of Example 1, b), “two-error channel” of Example 2, and c), the  $q$ -ary symmetric channel of Example 3.

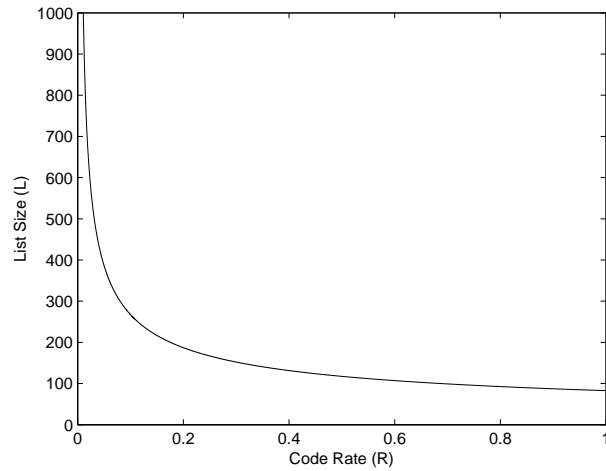


Fig. 2. List size using ASD for the “typewriter channel”.

Since the bound in (11) is a strictly decreasing function of the code's dimension  $k$ , an upper bound on the size of the list is obtained by taking  $k = 2$ :

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

Thus, the number of codewords on the list is bounded above by a slowly growing function of  $n$ .

#### D. A Closer Look at the ASD Error Radius

We are interested in quantifying when the error radius of ASD exceeds the error radius of GS decoding. For simplicity we only analyze the case of GS decoding with  $m \rightarrow \infty$  (note that it will also imply that ASD is better than GS decoding for any finite  $m$ ).

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

$$R < \left( \frac{p_{\min} - 1/\lambda}{\sqrt{\gamma + 1/\lambda} - p_{\max} + p_{\min}} \right)^2. \quad (12)$$

*Proof:* Solving the equation

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

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

Corollary 1 gives a sufficient condition for the ASD correction radius to exceed the GS decoding radius. This condition is nontrivial for a subset of low code rates and  $\lambda$  large enough (see 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 2 quantifies that region.

*Corollary 2:* Let  $\lambda > 1/p_{\min}$ . If

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

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

*Proof:* If the right-hand side of (5) is 1, then ASD will produce a list that contains  $\mathbf{c}$  regardless of the error pattern. Thus, the claim reduces to solving for  $R$  the inequality

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

from which the corollary follows.  $\square$

It is a surprising result that there exists non-zero rates where ASD always produces a list polynomial in  $n$  that contains the transmitted codeword. The intuition is that  $Pr(\mathbf{y}|\mathbf{c}) \neq 0$  for all channels, but for many codewords  $\mathbf{c}' \in C$ ,  $Pr(\mathbf{y}|\mathbf{c}') = 0$  due to zeros in the transition probability matrix. The zeros in the transition probability matrix allow the soft-decision decoder to discount codewords. When the code rate is low enough, the decoder can reduce the size of the list to be within the bound of Proposition 1 without actually eliminating possible codewords. In this case, the probability of list-decoding error is zero.

For the  $q$ -ary symmetric channel, a case where intuition tells us that ASD should provide no improvement over GS decoding, Figure 1(c) 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 condition  $\lambda > 1/p_{\min}$  and (12) for the  $q$ -ary symmetric channel unless we have  $\lambda \rightarrow \infty$ . 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.

#### IV. ASD PROBABILITY OF ERROR BOUNDS

This section is focused on bounding the probability of error for ASD. 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 the final stage of decoding, it may be required to select a unique codeword candidate from the list obtained using the maximum likelihood criterion. In this section, we derive a bound for the probability of error when the correct decoding corresponds to the case that the transmitted codeword is on the decoder's list and it is selected as the best estimate.

The only previous work that deals with deriving bounds on the probability of error for ASD is [15]. In that paper, Ratnakar and Koetter consider a general channel and use list-decoding error as the error event. We refine the results of [15] in two ways: first, we prove a tighter bound on the list-decoding error probability for low rates, and secondly, we derive a bound on the probability of selection error that is the dominating error event for those rates.

##### A. General Form

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

$$\mathcal{A} : \mathbf{c} \notin \mathcal{L}, \quad \mathcal{B} : \hat{\mathbf{c}} \neq \mathbf{c}.$$

The list-decoding probability of error, which is the probability that the list produced by ASD contains the transmitted codeword, is given by  $P\{\mathcal{A}\}$ . The selection probability of error is given by  $P\{\mathcal{B}\}$ . Observe that  $\mathcal{A} \subseteq \mathcal{B}$ , giving us  $P\{\mathcal{A}\} \leq P\{\mathcal{B}\}$ . Each of these probabilities can be bounded above by using the Chernoff bound (see e.g. [4]).

*Lemma 4: (Chernoff bound)* Let  $w$  be a random variable with moment generating function  $\Phi_w(s)$  and let  $A$  be a real number. Then

$$Pr\{w \geq A\} \leq e^{-sA} \Phi_w(s), \quad s > 0 \quad (13)$$

$$Pr\{w \leq A\} \leq e^{sA} \Phi_w(-s), \quad s > 0. \quad (14)$$

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

$$P\{\mathcal{A}\} \leq e^{-nE_A}, \quad P\{\mathcal{B}\} \leq e^{-nE_B}.$$

The functions  $E_A$  and  $E_B$  are the error exponents for  $P\{\mathcal{A}\}$  and  $P\{\mathcal{B}\}$ , respectively.

### B. List-Decoding Probability of Error

The next theorem gives an upper bound on the probability that the transmitted codeword is not on the list output by the decoder. A similar estimate of the error exponent, in a more general context, is derived in [15]. However [15] does not contain the analysis of the error radius that leads us to conclude that for low rates the transmitted codeword will always be on the list. In other words, the error event for these (and some other) rates will be dominated by an incorrect selection from the list obtained through ASD.

*Theorem 3:* The probability of event  $\mathcal{A}$  can be bounded above as

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

where

$$E_{\mathcal{A}} = -\ln \sum_{i=1}^q p_i e^{-s(p_i - \sqrt{R(\gamma+1/\lambda)} - 1/\lambda)}$$

except when  $R < \frac{(p_{\min} - 1/\lambda)^2}{\gamma + 1/\lambda}$  and  $\lambda > 1/p_{\min}$ , in which case  $E_{\mathcal{A}} = \infty$ .

*Proof:* The rate region where  $E_{\mathcal{A}} = \infty$  follows directly from Corollary 1. In order to gain insight into the event  $\mathcal{A}$  for the remainder of the rates, consider again the threshold condition for ASD list-decoding success given in Lemma 2. If the condition in Lemma 2 is met, it is clear that  $\mathcal{A}$  is false. However, if the condition is not met,  $\mathcal{A}$  could either be true or false. Thus, for a given  $\mathcal{M}$ , we have

$$P\{\mathcal{A}\} \leq \Pr\{S_{\mathcal{M}}(\mathbf{c}) < \sqrt{2(k-1)\mathcal{C}(\mathcal{M})}\}.$$

Assume now that the received vector  $\mathbf{y}$  is a random vector with coordinates distributed according to the transition probabilities in the channel. As above, let  $W_i$  be a random transition probability  $w_{c(i), y(i)}$ . With these assumptions, the components of the matrix of multiplicities as well as the score of a codeword  $\mathbf{c}$  and the cost of  $\mathcal{M}$  become random variables. Then

$$\Pr\{S_{\mathcal{M}}(\mathbf{c}) < \sqrt{2(k-1)\mathcal{C}(\mathcal{M})}\} \leq \Pr\left\{\sum_{i=1}^n \mathbf{W}_i < n\sqrt{R\left(\gamma + \frac{1}{\lambda}\right)} + \frac{n}{\lambda}\right\}.$$

Finally, the Chernoff bound (14) can be applied to give the result

$$P\{\mathcal{A}\} \leq e^{sn(\sqrt{R(\gamma+1/\lambda)} + \frac{1}{\lambda})} \left(\sum_{i=1}^q p_i e^{-sp_i}\right)^n$$

where  $s > 0$ . The theorem follows.  $\square$

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 no conclusion can be drawn regarding the possibility of reliable communication. 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  $\gamma$ .

*Lemma 5:*

$$E_{\mathcal{A}} > 0 \quad \text{if } R < \frac{(\gamma - 1/\lambda)^2}{\gamma + 1/\lambda}$$

*Proof:* We have that

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

First observe that  $g''(s) > 0$ , indicating that the function  $g(s)$  is convex. Next observe that  $g(0) = 1$  and that  $g'(0) < 0$  if and only if  $R < (\gamma - 1/\lambda)^2/(\gamma + 1/\lambda)$ . 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 gives a trivial estimate  $E_{\mathcal{A}} = 0$ .  $\square$

### C. Selection Probability of Error

When the list-decoding probability of error is zero as for the rate region defined in Corollary 2, one does not have insight into the performance of ASD. Therefore, it is of interest to quantify a comprehensive probability of error given by  $P\{\mathcal{B}\}$ . To make the selection error probability amenable to analysis we consider the ensemble of random Generalized Reed-Solomon (GRS) codes.

Let  $C$  be an RS code and  $\mathbf{w} = (w_1, \dots, w_n)$  be a vector of nonzero elements of  $\mathbb{F}_q$ . A  $\text{GRS}_k(\mathbf{w})$  code is obtained from  $C$  by multiplying every code vector  $\mathbf{c} \in C$  by a diagonal matrix  $\text{diag}(w_1, \dots, w_n)$ . Thus, the code  $C$  gives rise to the ensemble of  $(q-1)^n$   $\text{GRS}_k$  codes with uniform distribution on it induced by the choice of the modifying vector  $\mathbf{w}$ .

We will assume transmission with a random GRS code where  $\mathbf{w}$  is chosen uniformly before each transmission.

*Theorem 4:* Suppose that a random GRS code  $G$  of rate  $R$  is used to transmit through an additive channel. The decoding error probability of ASD can be bounded above as

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

where

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

*Proof:* Let  $\{c_1, \dots, c_M\}$  be the codewords of the code  $G$ . Define the events  $\mathcal{C}_i$  and  $\mathcal{D}_i$  as follows:

$$\mathcal{C}_i = \{(c_i \in \mathcal{L}) \& (S_{\mathcal{M}}(c_i) \geq S_{\mathcal{M}}(\mathbf{c}))\}$$

$$\mathcal{D}_i = \{S_{\mathcal{M}}(\mathbf{c}_i) \geq S_{\mathcal{M}}(\mathbf{c})\}$$

where  $\mathcal{L}$  is the decoder's list of codewords,  $\mathbf{c}$  is the transmitted codeword, and  $\mathcal{M}$  is the multiplicity matrix.

Observe that

$$P\{\mathcal{B}\} = \Pr\{\exists \mathbf{c}' \neq \mathbf{c}, \mathbf{c}' \in \mathcal{L} : S_{\mathcal{M}}(\mathbf{c}') \geq S_{\mathcal{M}}(\mathbf{c})\}$$

$$\begin{aligned} &= P\left\{ \bigcup_{\substack{i=1 \\ i: \mathbf{c}_i \neq \mathbf{c}}}^{q^k} \mathcal{C}_i \right\} \leq \sum_{\substack{i=1 \\ i: \mathbf{c}_i \neq \mathbf{c}}}^{q^k} P\{\mathcal{C}_i\} \\ &\leq \sum_{\substack{i=1 \\ i: \mathbf{c}_i \neq \mathbf{c}}}^{q^k} P\{\mathcal{D}_i\}. \end{aligned}$$

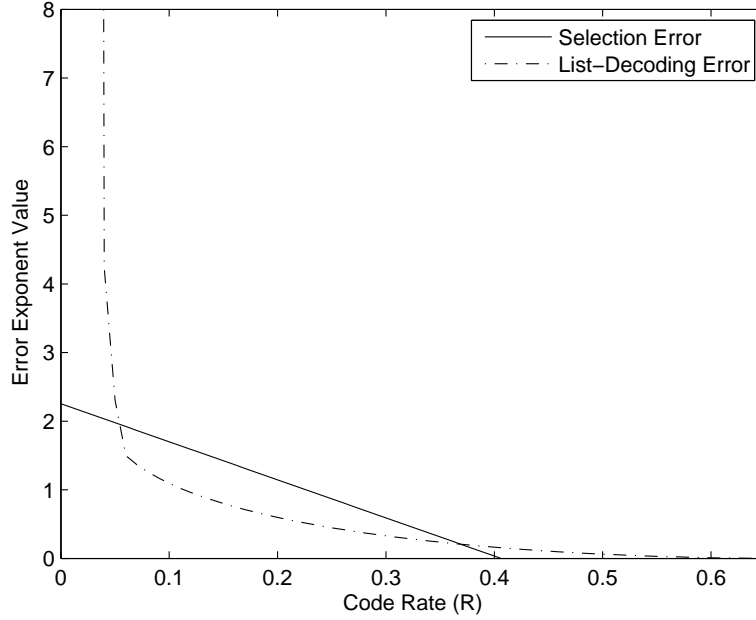


Fig. 3. Comparison of the two error exponents for ASD over low rates.

For all  $\mathbf{v}_i \in \mathbb{F}_q^n$ , define the event  $\mathcal{E}_i$  as follows:

$$\mathcal{E}_i : \{S_{\mathcal{M}}(\mathbf{v}_i) \geq S_{\mathcal{M}}(\mathbf{c})\}.$$

We have

$$\sum_{\substack{i=1 \\ i:\mathbf{c}_i \neq \mathbf{c}}}^{q^k} P\{\mathcal{D}_i\} = \sum_{\substack{i=1 \\ i:\mathbf{v}_i \neq \mathbf{c}}}^{q^n} P\{\mathcal{E}_i, \mathbf{v}_i \in G\}.$$

Further, every coordinate of  $\mathbf{c}$  is distributed uniformly in  $\mathbb{F}_q$  and therefore, the multiplicity  $m_{v(i),i}$  is a uniform random variable taking values in  $\{p_1, \dots, p_q\}$ . Moreover,  $\Pr(\mathbf{v}_i \in G) = q^{k-n}$ .

We then know that  $S_{\mathcal{M}}(\mathbf{v})$  is a sum of i.i.d. random variables  $m_{v(i),i}$ . As a result,  $P\{\mathcal{E}_i\}$  is the same for all  $\mathbf{v}_i \neq \mathbf{c}$ , and the events  $\mathcal{E}_i$  and  $\{\mathbf{v}_i \in C\}$  are independent, i.e.

$$\begin{aligned} \sum_{\substack{i=1 \\ i:\mathbf{v}_i \neq \mathbf{c}}}^{q^n} P\{\mathcal{E}_i, \mathbf{v}_i \in C\} &= \sum_{\substack{i=1 \\ i:\mathbf{v}_i \neq \mathbf{c}}}^{q^n} P\{\mathcal{E}_i\} P\{\mathbf{v}_i \in C\} \\ &= \sum_{\substack{i=1 \\ i:\mathbf{v}_i \neq \mathbf{c}}}^{q^n} \frac{1}{q^{n-k}} P\{\mathcal{E}_i\} \leq q^k P\{\mathcal{E}_i | \mathbf{v}_i \neq \mathbf{c}\}. \end{aligned}$$

Define the random variable  $W_i$  as in the proof of Theorem 1 and let  $V_i, i = 1, \dots, q^n$  be i.i.d. r.v.'s with

$$P(V_i = p_j) = 1/q, 1 \leq j \leq q.$$

Then

$$P\{\mathcal{B}\} \leq q^k P\{\mathcal{E}_i | \mathbf{v}_i \neq \mathbf{c}\}$$

$$\begin{aligned} &= q^k \Pr \left\{ \sum_{i=1}^n \lfloor \lambda V_i \rfloor \geq \sum_{i=1}^n \lfloor \lambda W_i \rfloor \right\} \\ &\leq q^k \Pr \left\{ \sum_{i=1}^n (V_i - W_i) \geq -\frac{n}{\lambda} \right\}. \end{aligned}$$

Now let us apply the Chernoff bound (13). For any  $s > 0$ ,

$$\begin{aligned} \Pr \left\{ \sum_{i=1}^n (V_i - W_i) \geq -\frac{n}{\lambda} \right\} \\ \leq e^{\frac{sn}{\lambda}} \left( \sum_{i=1}^q \frac{1}{q} e^{sp_i} \right)^n \left( \sum_{i=1}^q p_i e^{-sp_i} \right)^n. \end{aligned}$$

Algebraic manipulations and the appropriate definition of  $E_{\mathcal{B}}$  yield the theorem.  $\square$

It is of interest to find out the behavior of  $E_{\mathcal{A}}$  compared to  $E_{\mathcal{B}}$  across all values of  $R$ . Let us compare the two error exponents in an example.

*Example 5:* Consider once again Example 1 ( $p_{\max} = 0.8$ ,  $p_{\min} = 0.2$ ,  $\lambda = 100$ ,  $q = 256$ , and  $\gamma = 0.68$ ). Figure 3 shows the comparison of  $E_{\mathcal{A}}$  to  $E_{\mathcal{B}}$ . We 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.

## V. MULTIVARIATE INTERPOLATION

In this section, we estimate the decoding radius of ASD for a new class of codes introduced recently by Parvaresh and Vardy in [14]. These codes are constructed as evaluations of  $M \geq 2$  polynomials. A multivariate interpolation decoding algorithm for the codes in [14] is shown to exceed the GS decoding radius for low values of the code rate  $R$ . In this section we extend our analysis of ASD to multivariate interpolation.



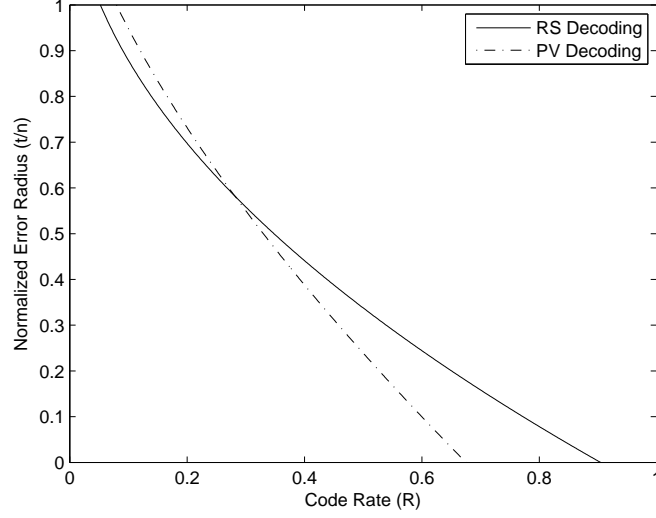


Fig. 4. Trivariate decoding of a PV code compared to ASD of a RS code.

A code  $C$  in the Parvaresh-Vardy (PV) family is defined as follows. Let  $\{1, \beta_1, \beta_2, \dots, \beta_{M-1}\}$  be a basis of  $\mathbb{F}_{q^M}$  over  $\mathbb{F}_q$ , let  $\{a_1, a_2, \dots, a_{M-1}\}$  be a set of positive integers greater than 1, and let  $e(X)$  be an irreducible polynomial over  $\mathbb{F}_q$ . Given a message  $\mathbf{u}$ , the encoder constructs  $f(X)$  as the polynomial derived from it and finds the set of polynomials  $\{g_1(X), g_2(X), \dots, g_{M-1}(X)\}$  by computing

$$g_i(X) = (f(X))^{a_i} \bmod e(X). \quad (15)$$

A codeword  $\mathbf{c} = \{c_1, c_2, \dots, c_n\}$  of a PV code that is associated with  $\mathbf{u}$  is found through the evaluation

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

It follows that the rate of a PV code  $C$  is  $R = k/Mn$  and the minimum distance is  $d = n - k + 1$ . Since (15) is a non-linear operation, the code  $C$  is not necessarily linear.

#### A. Soft-Decision Decoding of PV Codes

Although Parvaresh and Vardy only considered hard-decision decoding in [14], soft-decision decoding of Folded RS Codes, a broader class of codes that contains PV codes, was considered by Guruswami and Rudra in [6]. Remark 4 of [6] includes a condition for list-decoding success. In the next theorem we present a more stringent condition for list-decoding success of a PV code transmission.

**Theorem 5:** 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  $\mathbf{c} = (c_1, \dots, c_n)$  will be included in the list output by the algorithm if

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

where  $m_{i,j}$  is an element of the  $q^M \times n$  multiplicity matrix  $\mathcal{M}$ .

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

$$\text{wdeg } Q(X, Y_1, \dots, Y_M) < \left[ \sqrt[M+1]{(k-1)^M \sum_{i,j} \prod_{l=0}^M (m_{i,j} + l)} \right]. \quad (16)$$

where  $\text{wdeg } X^i Y_1^{j_1} \dots Y_M^{j_M} = i + (k-1) \sum_{l=1}^M j_l$ . If the score  $S_{\mathcal{M}}(\mathbf{c})$  exceeds the RHS of (16), then  $\mathbf{c}$  is on the algebraic soft-decision decoder's list by an argument similar to the one employed to prove Lemma 2.  $\square$

#### B. Multivariate Error Decoding Radius

Suppose the PV code is transmitted over a channel with transition probabilities  $\{p_1, p_2, \dots, p_{q^M}\}$ . The statistics  $p_{\min}$ ,  $p_{\max}$ , and  $\gamma$  are defined as before over this new set of transition probabilities. An error radius is given in Theorem 6 for soft-decision decoding of PV codes.

**Theorem 6:** Given a PV code with rate  $R = k/Mn$  is used to communicate over an additive-noise channel. If

$$\frac{t}{n} \leq \frac{p_{\max} - \sqrt[M+1]{\frac{R^M M^M}{(M+1)!} \sum_{i=1}^{q^M} \prod_{l=0}^M (p_i + l/\lambda)} - \frac{1}{\lambda}}{p_{\max} - p_{\min}}, \quad (17)$$

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

The proof is very similar to the proof of Theorem 1 and is omitted. The multivariate ASD error radius is larger than the bivariate ASD error radius for low-rate codes. This claim is shown through an example.

**Example 6:** 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 (in other words, ASD). Figure 4 shows the error radii (5) and (17). The graph shows that trivariate decoding provides an improvement over the bivariate one for rates less than 0.3.

## VI. CONCLUSION

The results presented in this paper have shown that soft-decision algebraic list decoding of RS and related codes is able to outperform its hard-decision counterparts for low-rate to medium-rate codes. An estimate of the error decoding radius derived in the paper enables ASD to be compared for the first time to other RS decoding methods. This result has also been extended to multivariable RS codes. A better estimate of the error probability for list decoding under ASD is derived which shows that at lower rates, the list decoding error is not an adequate performance criterion for this algorithm. A comprehensive probability of error bound is also derived that includes the previously overlooked probability of selection error.

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, an interesting (and important in applications) fact of ASD decoding outperforming its hard-decision counterparts for high-rate codes claimed in some experimental studies, so far has not been confirmed by theoretical analysis.

## REFERENCES

- [1] R. E. Blahut, *Theory and Practice of Error-Correcting Codes*, Reading, MA: Addison-Wesley, 1983.
- [2] T. Cover and J. Thomas, *Elements of Information Theory*, New York, N.Y.: 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," in: A. Ashikhmin and A. Barg (Eds.), *Algebraic Coding Theory and Information Theory*, pp. 99-120, Providence, RI: AMS, 2005.
- [4] R. Gallager, *Information Theory and Reliable Communication*, New York, N.Y.: John Wiley and Sons, 1968.
- [5] W. Gross, F. Kschischang, R. Koetter, and G. Gulak, "Applications of algebraic soft-decision decoding of Reed-Solomon codes," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1224-1234, 2006.
- [6] V. Guruswami and A. Rudra, "Explicit capacity-achieving list-decodable codes," Electronic Colloquium on Computational Complexity, Report 05-133, 2005, online at <http://eccc.hpi-web.de/eccc/>.
- [7] V. Guruswami and M. Sudan, "Improved decoding of Reed-Solomon and algebraic-geometric codes," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1757-1767, 1999.
- [8] J. Jiang and K. Narayanan, "Performance analysis of algebraic soft decoding of Reed-Solomon codes over binary symmetric and erasure channels," *Proc. 2005 IEEE Internat. Sympos. Inform. Theory*, Adelaide, Australia, Sept. 4-9, p. 1186-1190, 2005.
- [9] J. Justesen, "Soft-decision decoding of RS codes," *Proc. 2005 IEEE Internat. Sympos. Inform. Theory*, Adelaide, Australia, Sept. 4-9, 2005, pp. 1183-1185.
- [10] R. Koetter, "On optimal weight assignments for multivariate interpolation list-decoding," *Proc. 2006 IEEE Information Theory Workshop*, Punta del Este, Uruguay, March 13-17, pp. 37-41, 2006.
- [11] R. Koetter and A. Vardy, "Algebraic soft-decision decoding of Reed-Solomon codes," *IEEE Trans. Inform. Theory*, vol. 49, no. 11, pp. 2809-2825, 2003.
- [12] R. McEliece, "The Guruswami-Sudan decoding algorithm for Reed-Solomon Codes," manuscript, 2003, available on-line at <http://www.systems.caltech.edu/EE/Faculty/rjm/>.
- [13] F. Parvaresh and A. Vardy, "Multiplicity assignments for algebraic soft-decision decoding of Reed-Solomon Codes," *Proc. 2005 IEEE Internat. Sympos. Inform. Theory*, Yokohama, Japan, June 29-July 3, 2003, p. 250.
- [14] F. Parvaresh and A. Vardy, "Correcting errors beyond the Guruswami-Sudan radius in polynomial time," *Proc. 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. 51, no. 11 pp. 3899-3917, 2005.