On the Multiuser Error Probability and the Maximum Likelihood Performance of MDS Codes.

Mostafa El-Khamy * and Robert J. McEliece**

Electrical Engineering Department

California Institute of Technology

Pasadena, CA 91125, USA

*E-mail: mostafa@systems.caltech.edu **E-mail: rjm@systems.caltech.edu

Abstract

MDS (e.g. Reed-Solomon) codes have many desirable properties which make them the code of choice in network scenarios and distributed coding schemes. An average binary weight enumerator of Reed-Solomon (RS) codes is derived assuming a binomial distribution of the bits in a non-zero symbol. Lower bounds on the average binary minimum distance of the ensemble of binary images of a Reed-Solomon code are shown. The ensemble of binary images of the RS code is shown to be, on average, asymptotically good. The performance of bit-level Reed-Solomon maximum likelihood decoders is studied. Given an arbitrary partition of the coordinates of a code, we introduce the partition weight enumerator which enumerates the codewords with a certain weight profile in the partitions. A closed form formula of the partition weight enumerator of maximum distance separable (MDS) codes is derived. Using this result, some properties of MDS codes are discussed. In particular, we show that all coordinates have the same weight within the subcodes of constant weight codewords. The results are extended for the ensemble of binary images of MDS codes defined over finite fields of characteristic two. The error probability of Reed-Solomon codes in multiuser networks is then studied.

This research was supported by NSF grant no. CCF-0514881 and grants from Sony, Qualcomm, and the Lee Center for Advanced Networking. This work was presented in part at the 2004 42nd Allerton Conf. on Communication, Control and Computing [1] and at the 2005 IEEE International Symposium on Information Theory, Adelaide, Australia [2].

I. Introduction and Outline

Reed-Solomon (RS) codes are the most popular maximum distance separable (MDS) codes. For any linear (n, k, d) code (of length n, dimension k and minimum distance d) over any field, maximum distance separable (MDS) codes have the maximum possible minimum distance $d = n - k + 1$ [3]. MDS codes have many other desirable properties which made them the code of choice in many communication systems. MDS codes have the property that any k codeword coordinates can be considered as the information symbols in a systematic codeword and any k coordinates can be used to recover the information symbols. Moreover, punctured MDS codes are also MDS codes. Such properties made MDS codes a natural choice in Automatic-Repeat-Request (ARQ) communication systems (c.f. [4]). MDS codes are also used in the design of multicast network codes [5]. Reed Solomon codes are one of the most important linear block codes and have been deployed in a wide range of applications [6]. Maximum-likelihood decoding of linear block codes is well-known to be NP-hard [7]. The Guruswami-Sudan (GS) algorithm was the first polynomial time hard-decision decoding algorithm for Reed-Solomon codes capable of correcting beyond half-the-minimum distance of the code [8]. Moreover, the invention of the GS algorithm has spurred a significant amount of research aiming at better soft-decision decoding algorithms for Reed-Solomon codes (c.f. [9], [10], [11], [12], [13]).

Suppose a Reed-Solomon (RS) code is defined over a finite field of characteristic two, then it is a common practice to send its binary image over the channel. In fact, the binary image has a large burst-error correction capability which is one of the main reasons behind the ubiquitous use of RS codes. The decoder can either be a bit-level decoder, which decodes the RS code as a binary code, or a symbol level decoder, which treats the received word as a vector in the finite field. It is often the case that hard-decision decoders, which do not make use of the reliability information from the channel, are symbol based decoders. Such hard-decision decoders, as the Berlekamp-Massey algorithm and the Guruswami-Sudan algorithm, usually operate on the symbol level to make use of the nice algebraic properties of RS codes. Soft-decision decoders make use of the channel reliability information. In case the code is sent over a binary input channel, then the decoder is often a bit-level decoder. With the recent advances in soft-decision decoding of RS codes, it was vital to benchmark the performance of such algorithms against the optimum soft-decision maximum likelihood decoder.

A significant amount of research has been recently devoted to finding tight bounds on the performance of linear codes under maximum-likelihood decoding [14]. The maximum-likelihood performance of linear codes requires the knowledge of the weight enumerator. Unfortunately, knowing the weight enumerator of the binary images of RS codes is very hard. Some attempts have been successful in giving the binary weight enumerator for particular realizations of RS codes [15]. Other researchers considered enumerating the codewords by the number of symbols of each kind in each codeword [16]. The average binary weight enumerators of a class of generalized Reed-Solomon codes, derived from an original RS code either by using a different basis to expand each column in the RS generator matrix into a binary representation or by multiplying each column in the RS generator matrix by some non-zero element in the field, were also studied [17].

Consider a network scenario, where users in a certain cluster can communicate in an error free manner. These users would like to communicate with another set of users in another cluster over a noisy channel. If the users in the first cluster are of limited power they will not be able to reliably transmit their information to the users in the other cluster. One solution is for the users to transmit their information to a local base station, which will then group their data symbols, encode them with a channel code and transmit the codeword to the other set of users. (See Fig. 1.) In other words, each codeword will be partitioned among more than one user or application. After decoding at the receiving base station, the information will be routed to the desired users. One other advantage of sharing a codeword among different users is the expected improvement in the code performance as its length increases [18]. Moreover, the recent results on the capacity of wireless networks suggest that networks with a smaller number of users and clustered networks are more likely to find acceptance [19]. Using the results in this paper, we will be able to analyze the performance of different users in such a scenario.

This paper is organized as follows:

In Section II, we introduce a generalized weight enumerator, which we call the partition weight enumerator (PWE). Given a partition of the coordinates of a code, the PWE enumerates the codewords with a certain weight profile in the partitions. Our main result is a simple closed-form expression for the PWE of an arbitrary MDS, e.g., Reed-Solomon, code (Section III, Theorem 6). This generalizes the results of Kasami et al. [20] on the split weight enumerator of RS codes. The PWE is a very useful tool in proving some of the nice algebraic properties of

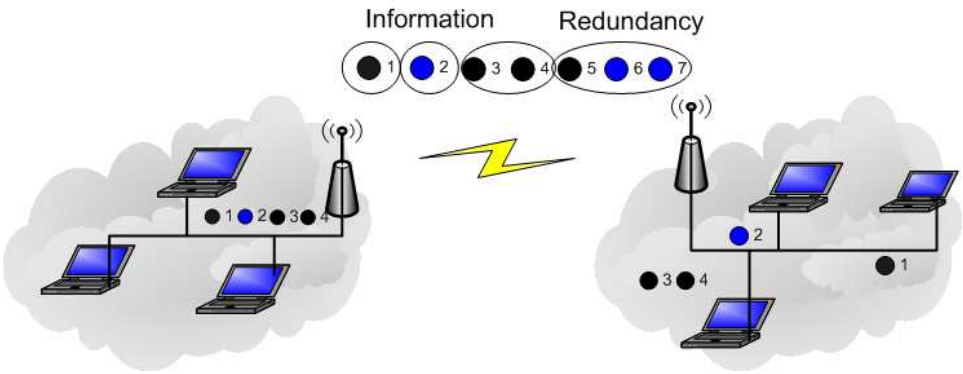

Fig. 1. A multiuser scenario where users within the same cluster transmit their information to a local base station, which, in turn, groups their symbols into one data word and transmits it, after channel encoding, over a noisy channel to the users in another cluster.

MDS codes. We then proceed in Section IV to derive a strong symmetry property for MDS codes (Theorem 8) which allows us to obtain improved bounds on the symbol error probability for RS codes. We show that an approximation widely used to estimate the symbol error probability of linear codes is exact for MDS codes. We take this opportunity to discuss other codes which also have this property.

One of the main motivations behind this paper was the following question:
How can one analyze the maximum-likelihood performance of the binary images of RS codes? In Section V, we attempt to answer this question by studying the weight enumerator of the ensemble of binary images of Reed Solomon codes. In fact we show that the ensemble weight enumerator approaches that of a random code with the same dimension. It is also well known that the minimum distance of a linear code provides a lot of insight about its performance. This motivated us to study the minimum distance of the ensemble of binary images of RS codes (Section VI). We show that the ensemble has an asymptotically good minimum distance. Given this result, one can search for good codes within the ensemble of binary images of Reed Solomon codes. We then attempt to answer the above question in Section VII, where we analyze the performance of soft and hard-decision maximum likelihood decoding of the binary images of the RS code. We show that the bounds developed using the techniques in this paper are indeed tight.

As we have mentioned, the ensemble average weight enumerators of the binary images of RS

codes have been rendered useful in analyzing their performance. We also study the case when the binary images of an Reed-Solomon is partitioned among different users or applications. In Section VIII, we show that the ensemble also has a similar symmetry property which becomes useful when analyzing its bit error probability.

As an application to the results in this paper, we study, in Section IX, the codeword, symbol and bit error probabilities of various Reed-Solomon code decoders in a generalized setting. In Section X, we prove that if systematic MDS (e.g. RS) codes are used in a multiuser setting, the unconditional symbol or bit error probabilities of all the users will be the same regardless of the size of the partitions assigned to them. We also considered various network scenarios where the Reed-Solomon code is the channel code of choice. We also proceed to show how one can analyze the error probability of a certain user given some conditions on the performance of other users. In Section XI, we conclude the paper and give some insights about the results in this paper.

II. Weight Enumerators

We begin by generalizing the notion of Hamming weight. Let $\mathbb{F}_q^n$ denote the vectors of length n over the finite field of q elements $\mathbb{F}_q$. A linear code $\mathcal{C}$ of length n defined over $\mathbb{F}_q$ is a linear subspace of $\mathbb{F}_q^n$. Let $N = \{1, 2, \ldots, n\}$ be the coordinate set of $\mathcal{C}$. Suppose N is partitioned into p disjoint subsets $N_1, \ldots, N_p$, with $|N_i| = n_i$, for $i = 1, \ldots, p$ [1]. We stress that $\sum_{i=1}^{p} n_i = n$. The elements of the set $N_i \subset N$ are given by $N_i = \{N_i(1), N_i(2), ..., N_i(n_i)\}$. Let $\boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n)$ be a vector in $\mathbb{F}_q^n$, then the ith partition of $\boldsymbol{v}$ is the vector $\boldsymbol{v}[N_i] = \big(\boldsymbol{v}_{N_i(1)}, \boldsymbol{v}_{N_i(2)}, ..., \boldsymbol{v}_{N_i(n_i)}\big)$.

Note that the number of ways a set of n coordinates could be partitioned into m_1 partitions of size of p_1, m_2 partitions of size p_2 and m_r of size p_r, i.e. the total number of partitions is $\sum_{i=1}^{r} m_r$ and $n = \sum_{i=1}^{r} m_r p_r$), is

$$\frac{n!}{\prod_{i=1}^{r} (p_i!)^{m_i} m_i!}, \tag{1}$$

where $x!$ is the factorial of x and the multinomial coefficient is normalized by the factor $\prod_{i=1}^{r} m_i!$ as we do not distinguish between partitions of the same size.

Denoting an $(n_1, ..., n_p)$ partition by $\mathcal{T}$, the $\mathcal{T}$-weight profile of a vector $\boldsymbol{v} \in \mathbb{F}_q^n$ is defined as $\mathcal{W}_{\mathcal{T}}(\boldsymbol{v}) = (w_1, \ldots, w_p)$, where w_i is the Hamming weight of $\boldsymbol{v}$ restricted to N_i, i.e., the

[1] Throughout this paper, the cardinality of a set T will be denoted by $|T|$.

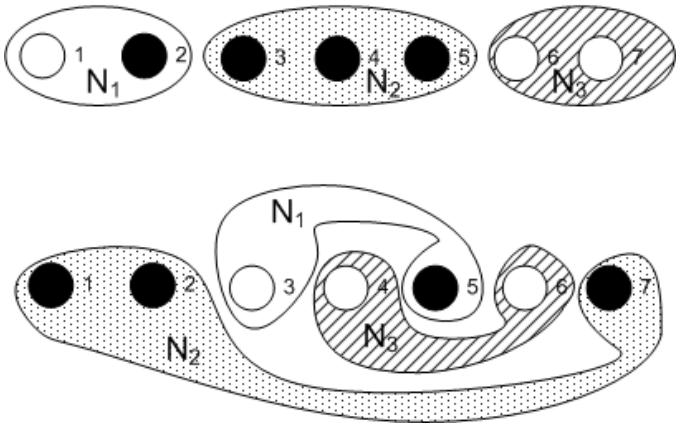

Fig. 2. The figure shows two different vectors in $\mathbb{F}_q^7$ and two different $\mathcal{T} : (2, 3, 2)$ partitions are applied. The weight profile of the vectors is $\mathcal{W}_{\mathcal{T}}(\boldsymbol{v}) = (1, 3, 0)$ where the zero and non-zero symbols are represented by white and black circles respectively.

weight of the vector $\boldsymbol{v}(N_i)$. (For an example see Fig. 2.) Given a code $\mathcal{C}$ of length n, the weight enumerator of $\mathcal{C}$ is

$$E_{\mathcal{C}}(w) = |\{\boldsymbol{c} \in \mathcal{C} : \mathcal{W}(\boldsymbol{c}) = w\}|, \tag{2}$$

where $\mathcal{W}(\boldsymbol{c})$ is the Hamming weight of $\boldsymbol{c}$. The weight generating function (WGF) of $\mathcal{C}$ is the polynomial

$$\mathbb{E}_{\mathcal{C}}(\mathcal{X}) = \sum_{h=0}^{n} E_{\mathcal{C}}(h)\mathcal{X}^h, \tag{3}$$

where the coefficient of $\mathcal{X}^h$ is the number of codewords with weight h;

$$E_{\mathcal{C}}(h) = \text{Coeff}\left(\mathbb{E}_{\mathcal{C}}(\mathcal{X}), \mathcal{X}^h\right). \tag{4}$$

(The subscript $\mathcal{C}$ may be dropped when there is no ambiguity about the code.) Now we generalize the notion of code weight enumerator. For an $(n_1, n_2, ..., n_p)$ partition $\mathcal{T}$ of the n coordinates of $\mathcal{C}$, the $\mathcal{T}$-weight enumerator of $\mathcal{C}$ enumerates the codewords with a weight profile $(w_1, \ldots, w_p)$

$$A_{\mathcal{C}}^{\mathcal{T}}(w_1, \ldots, w_p) = |\{\boldsymbol{c} \in \mathcal{C} : \mathcal{W}_{\mathcal{T}}(\boldsymbol{c}) = (w_1, \ldots, w_p)\}|.$$

The *partition weight generating function* (PWGF) is given by the multivariate polynomial

$$\mathbb{P}^{\mathcal{T}}(\mathcal{X}_1, ..., \mathcal{X}_p) = \sum_{w_1=0}^{n_1} ... \sum_{w_p=0}^{n_p} A^{\mathcal{T}}(w_1, ..., w_p)\mathcal{X}_1^{w_1}...\mathcal{X}_p^{w_p}. \tag{5}$$

For the special case of two partitions, ($p = 2$), $A^{\mathcal{T}}(w_1, w_2)$ is termed the *split weight enumerator* in the literature [3]. The *input-redundancy weight enumerator* (IRWE) $R(w_1, w_2)$ is

the number of codewords with input weight (weight of the information vector) w_1 and redundancy weight w_2. For a systematic code, if $\mathcal{T}$ is an $(k, n-k)$ partition such that the first partition constitutes of the coordinates of the information symbols, then $R(w_1, w_2) = A^{\mathcal{T}}(w_1, w_2)$. The *input-output weight enumerator* (IOWE) $O(w, h)$ enumerates the codewords of total Hamming weight h and input weight w. Assuming that the first partition constitutes of the information symbols, then $O(w, h) = R(w, h-w)$. For an $(k, n-k)$ partition $\mathcal{T}$, it is straight forward that

$$E(h) = \sum_{w=0}^{k} A^{\mathcal{T}}(w, h-w) = \sum_{w=0}^{k} O(w, h). \tag{6}$$

It is useful to know the IOWE and IRWE of a code when studying its bit error probability (e.g. [21]). The *input-output weight generating function*, $\mathbb{O}(\mathcal{X}, \mathcal{Y})$, and the *input-redundancy weight generating function*, $\mathbb{R}(\mathcal{X}, \mathcal{Y})$, of an (n, k) code are defined to be respectively,

$$\mathbb{O}(\mathcal{X}, \mathcal{Y}) = \sum_{w=0}^{k} \sum_{h=0}^{n} O(w, h) \mathcal{X}^w \mathcal{Y}^h, \tag{7}$$

$$\mathbb{R}(\mathcal{X}, \mathcal{Y}) = \sum_{w_1=0}^{k} \sum_{w_2=0}^{n-k} R(w_1, w_2) \mathcal{X}^{w_1} \mathcal{Y}^{w_2}. \tag{8}$$

Since every non-zero symbol in the redundancy part of the code contributes to both its output and redundancy weights, $\mathbb{R}(\mathcal{X}, \mathcal{Y})$ and $\mathbb{O}(\mathcal{X}, \mathcal{Y})$ are related by the following transformations

$$\mathbb{R}(\mathcal{X}, \mathcal{Y}) = \mathbb{O}\left(\frac{\mathcal{X}}{\mathcal{Y}}, \mathcal{Y}\right), \quad \mathbb{O}(\mathcal{X}, \mathcal{Y}) = \mathbb{R}\left(\mathcal{X}\mathcal{Y}, \mathcal{Y}\right), \quad \mathbb{E}(\mathcal{X}) = \mathbb{R}(\mathcal{X}, \mathcal{X}). \tag{9}$$

For a systematic code, let the jth partition constitute of information symbols, then the jth IOWE enumerates the codewords with a Hamming weight w in the jth partition and a total weight h,

$$O^j(w, h) = |\left\{\boldsymbol{c} \in \mathcal{C} : \left(\mathcal{W}\left(\boldsymbol{c}[N_j]\right) = w\right) \wedge \left(\mathcal{W}(\boldsymbol{c}) = h\right)\right\}|, \tag{10}$$

and is derived from the PWGF by

$$\mathbb{O}^j(\mathcal{X}, \mathcal{Y}) = \mathbb{P}^{\mathcal{T}}(\mathcal{Y}, \mathcal{Y}, ., \mathcal{X}\mathcal{Y}, ., \mathcal{Y}) = \sum_{w=0}^{n_j} \sum_{h=0}^{n} O^j(w, h) \mathcal{X}^w \mathcal{Y}^h \tag{11}$$

where the invariants $\mathcal{X}_i$s in $\mathbb{P}^{\mathcal{T}}_{\mathbb{C}}(\mathcal{X}_1, \mathcal{X}_2, ..., \mathcal{X}_p)$ are substituted by

$$\begin{cases} \mathcal{X}_i := \mathcal{Y}, & \forall \quad i \neq j \\ \mathcal{X}_i := \mathcal{X}\mathcal{Y}, & i = j. \end{cases} \tag{12}$$

III. Partition weight enumerator of MDS codes

For an (n,k,d) MDS code over $\mathbb{F}_q$, it is well know that the minimum distance is $d=n-k+1$ [22] and that the weight distribution is given by [23, Theorem 25.7]

$$E(i) = \binom{n}{i}\sum_{j=d}^{i}\binom{i}{j}(-1)^{i-j}(q^{j-d+1}-1) \quad (13)$$

$$= \binom{n}{i}(q-1)\sum_{j=0}^{i-d}(-1)^{j}\binom{i-1}{j}q^{i-j-d}, \quad (14)$$

for weights $i \geq d$. In the next theorem, we show that for an arbitrary partition of the coordinates of an MDS code, and for any number of partitions, the partition weight enumerator of MDS codes admits a closed form formula.

Theorem 1: For an (n,k,d) MDS code $\mathcal{C}$ defined over $\mathbb{F}_q$, let $\mathcal{T}$ define a p-partition of the coordinates of $\mathcal{C}$ into p mutually exclusive subsets, N_1, N_2,..., N_p, such that $N_1 \cup N_2 ... \cup N_p = N$ where $N=\{1,2,...,n\}$ and $|N_i|=n_i$. The p-partition weight enumerator is given by

$$\binom{n_1}{w_1}....\binom{n_p}{w_p}\sum_{j_1=0}^{w_1}\binom{w_1}{j_1}(-1)^{w_1-j_1}\sum_{j_2=0}^{w_2}\binom{w_2}{j_2}(-1)^{w_2-j_2}$$
$$.... \sum_{j_p=d-\sum_{z=1}^{p-1}j_z}^{w_p}\binom{w_p}{j_p}(-1)^{w_p-j_p}(q^{\sum_{z=1}^{p}j_z-d+1}-1).$$

Proof: For $i=1,2,..p$, let R_i be a subset of N_i. Define $S(\boldsymbol{c})$ to be the support set of the codeword $\boldsymbol{c}$, i.e. the set of indices of the non-zero elements. Define

$$f(R_1,R_2,..,R_p) \stackrel{\Delta}{=} |\boldsymbol{c}\in\mathcal{C} : \{S(\boldsymbol{c})\cap N_i\} = R_i \ \ \forall i| = |\boldsymbol{c}\in\mathcal{C} : \{S(\boldsymbol{c}) = \bigcup_{i=1}^{p} R_i\}| \quad (15)$$

to be the number of codewords which are exactly non-zero on the sets R_i. From the definition of the p-partition weight enumerator, it follows that

$$A^{\mathcal{T}}(w_1,w_2,...,w_p) = \sum_{\substack{R_1\subseteq N_1\\ |R_1|=w_1}}\sum_{\substack{R_2\subseteq N_2\\ |R_2|=w_2}} ... \sum_{\substack{R_p\subseteq N_p\\ |R_p|=w_p}} f(R_1,R_2,...,R_p). \quad (16)$$

Define the mutually exclusive subsets, $S_i \subseteq N_i$, $i=1,2,..,p$ and let

$$g(S_1,S_2,...,S_p) = \sum_{R_1\subseteq S_1}\sum_{R_2\subseteq S_2} ... \sum_{R_p\subseteq S_p} f(R_1,R_2,...,R_p) \quad (17)$$

to be the number of codewords which are always zero on the sets $N_i \setminus S_i$ (See Fig. 3.). It follows from the MDS property of the code that if only m symbols of an (n,k) MDS code are

allowed to be non-zero, the $n-m$ zero symbols could be taken as information symbols, then the dimension of the resulting subcode is $k-n+m$ and

$$g(S_1,S_2,...,S_p)=\begin{cases}1, & \sum_{i=1}^{p}|S_i|<d;\\ q^{1-d+\sum_{i=1}^{p}|S_i|}, & n\geq\sum_{i=1}^{p}|S_i|\geq d,\end{cases}, \tag{18}$$

Successively applying Möbius Inversion [23, Theorem 25.1] to (17), we get

$$\begin{aligned}f(R_1,R_2,...,R_p)&=\sum_{S_1\subseteq R_1}\mu(S_1,R_1)...\sum_{S_p\subseteq R_p}\mu(S_p,R_p)g(S_1,S_2,...,S_p)\\&\triangleq\prod_{i=1}^{p}\left(\sum_{S_i\subseteq R_i}\mu(S_i,R_i)\right)g(S_1,S_2,...,S_p),\end{aligned} \tag{19}$$

where

$$\mu(S,R)=\begin{cases}(-1)^{|R|-|S|}, & S\subseteq R;\\ 0, & \text{otherwise.}\end{cases} \tag{20}$$

Substituting (19) in (16),

$$\begin{aligned}A^{\mathcal{T}}(w_1,w_2,...,w_p)&=\prod_{i=1}^{p-1}\left(\sum_{\substack{R_i\subseteq N_i\\|R_i|=w_i}}\sum_{S_i\subseteq R_i}(-1)^{|R_i|-|S_i|}\right)G_p(\beta)\\&=\prod_{i=1}^{p-1}\left(\binom{n_i}{w_i}\sum_{j=0}^{w_i}\binom{w_i}{j}(-1)^{w_i-j}\right)G_p(\beta),\end{aligned} \tag{21}$$

such that $\beta=\sum_{i=1}^{p-1}|S_i|$ and by invoking (18)

$$\begin{aligned}G_p(\beta)&=\sum_{\substack{R_p\subseteq N_p\\|R_p|=w_p}}\sum_{S_p\subseteq R_p}(-1)^{|R_p|-|S_p|}g(S_1,S_2,...,S_p)\\&=\binom{n_p}{w_p}\left(\sum_{i=0}^{d-\beta-1}\binom{w_p}{i}(-1)^{w_p-i}+\sum_{i=d-\beta}^{w_p}\binom{w_p}{i}(-1)^{w_p-i}q^{i+\beta-d+1}\right)\\&=\binom{n_p}{w_p}\sum_{i=d-\beta}^{w_p}\binom{w_p}{i}(-1)^{w_p-i}(q^{i+\beta-d+1}-1)\end{aligned} \tag{22}$$

The last equality follows from the fact that $\sum_{j=0}^{w}\binom{w}{j}(-1)^{w-j}=(1-1)^w=0$. Substituting (19) in (16), the theorem follows. ■

For the special case of two partitions, the split weight enumerator $A_{w_1,w_2}(n_1,n_2)$ is given in the following corollary.

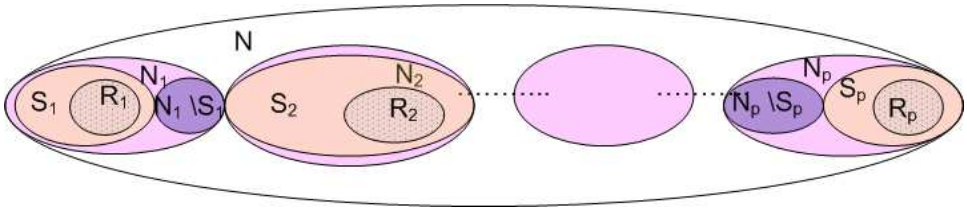

Fig. 3. The code is always zero on the coordinates in the sets $N_i \setminus S_i$ for $i = 1, 2, ..., p$.

Corollary 2: Let $\mathcal{T}$ be an (n_1, n_2) partition of an (n, k, d) MDS code $\mathcal{C}$, then the split weight enumerator of $\mathcal{C}$ is

$$A^{\mathcal{T}}(w_1, w_2) = \binom{n_1}{w_1}\binom{n_2}{w_2}\sum_{j=0}^{w_1}\binom{w_1}{j}(-1)^{w_1-j}\sum_{i=d-j}^{w_2}\binom{w_2}{i}(-1)^{w_2-i}(q^{i+j-d+1}-1).$$

From Theorem 1, it follows that the PWE of MDS codes does not depend on the orientation of the coordinates with respect to the partitions but only on the partitions' sizes and weights (see (17)). It thus intuitive that the ratio of $A^{\mathcal{T}}(w_1, w_2, ..., w_p)$ to $E(w)$ where $w = \sum_{i=1}^{p} w_i$ is the probability that the w nonzero symbols are distributed among the partitions with a $\mathcal{T}$-profile $(w_1, w_2, ..., w_p)$. Next we calculate this probability for the special case of $p = 2$ and we show that the partition weight enumerator admits to a simpler closed form formula.

Theorem 3: Let $\mathcal{T}$ be an (n_1, n_2) partition for an (n, k, d) MDS code, $n = n_1 + n_2$, then

$$A^{\mathcal{T}}(w_1, w_2) = E(w_1 + w_2)\frac{\binom{n_1}{w_1}\binom{n_2}{w_2}}{\binom{n}{w_1+w_2}}.$$

Proof: From Corollary 2, the split weight enumerator is

$$A^{\mathcal{T}}(w_1, w_2) = \binom{n_1}{w_1}\binom{n_2}{w_2}\sum_{j=0}^{w_1}\binom{w_1}{j}(-1)^{w_1-j}\sum_{i=d-j}^{w_2}\binom{w_2}{i}(-1)^{w_2-i}(q^{i+j-d+1}-1). \quad (23)$$

Doing a change of variables, $\alpha = i + j$, we get

$$A^{\mathcal{T}}(w_1, w_2) = \binom{n_1}{w_1}\binom{n_2}{w_2}\sum_{j=0}^{w_1}\binom{w_1}{j}(-1)^{w_1-j}\sum_{\alpha=\max(d,j)}^{w_2+j}\binom{w2}{\alpha-j}(-1)^{w_2-\alpha+j}(q^{\alpha-d+1}-1).$$

By changing the order of summation and summing over the same region:

$$\begin{aligned} A^{\mathcal{T}}(w_1, w_2) = \quad & \binom{n_1}{w_1}\binom{n_2}{w_2}\sum_{\alpha=d}^{w_1+w_2}(q^{\alpha-d+1}-1)(-1)^{w_1+w_2-\alpha}\sum_{j=0}^{\min(\alpha,w_1)}\binom{w_1}{j}\binom{w_2}{\alpha-j} \\ & -\binom{n_1}{w_1}\binom{n_2}{w_2}\sum_{\alpha=w_2+1}^{w_1+w_2}(q^{\alpha-d+1}-1)(-1)^{w_1+w_2-\alpha}\sum_{j=0}^{\alpha-w_2-1}\binom{w_1}{j}\binom{w_2}{\alpha-j} \end{aligned}$$

By doing the change of variables $\beta = \alpha - w_2$ in the second summation

$$A^{\mathcal{T}}(w_1, w_2) = \binom{n_1}{w_1}\binom{n_2}{w_2}\sum_{\alpha=d}^{w_1+w_2}(q^{\alpha-d+1}-1)(-1)^{w_1+w_2-\alpha}\binom{w_1+w_2}{\alpha}$$
$$-\binom{n_1}{w_1}\binom{n_2}{w_2}\sum_{\beta=1}^{w_1}(q^{\alpha-d+1}-1)(-1)^{w_1+w_2-\alpha}\sum_{j=0}^{\beta-1}\binom{w_1}{j}\binom{w_2}{w_2+\beta-j}.$$

Since $\beta - j$ is always positive it follows that the second term in the right hand side is always zero and by letting $w = w_1 + w_2$

$$A^{\mathcal{T}}(w_1, w_2) = \binom{n_1}{w_1}\binom{n_2}{w_2}\sum_{\alpha=d}^{w}\binom{w}{\alpha}(-1)^{w-\alpha}(q^{\alpha-d+1}-1). \tag{24}$$

By comparing with (13), the result follows. ■

Corollary 4: The IOWE of a systematic MDS code, $O(w,h)$, for $h \geq d$, is given by

$$O(w,h) = R(w,h-w) = E(h)\frac{\binom{k}{w}\binom{n-k}{h-w}}{\binom{n}{h}}$$
$$= \binom{k}{w}\binom{n-k}{h-w}\sum_{j=0}^{w}\binom{w}{j}(-1)^{w-j}\sum_{i=d-j}^{h-w}\binom{h-w}{i}(-1)^{h-w-i}(q^{i+j-d+1}-1).$$

By observing (6) and defining $\Psi(w)$ to be

$$\Psi(w) = \sum_{j=0}^{w}\binom{w}{j}(-1)^{w-j}\sum_{i=d-j}^{h-w}\binom{h-w}{i}(-1)^{h-w-i}(q^{i+j-d+1}-1), \tag{25}$$

we have an interesting identity:

$$\sum_{w=0}^{k}\Psi(w)\binom{k}{w}\binom{n-k}{h-w} = \Psi(0)\sum_{w=0}^{k}\binom{k}{w}\binom{n-k}{h-w}, \tag{26}$$

where $\binom{n}{h} = \sum_{w=0}^{k}\binom{k}{w}\binom{n-k}{h-w}$ and $\Psi(0) = \sum_{i=d}^{h}\binom{h}{i}(-1)^{h-i}(q^{i-d+1}-1)$.

Corollary 5: For an (n,k,d) MDS code $\mathcal{C}$, the number of codewords which are exactly nonzero at a fixed subset of coordinates of cardinality h and are zero at the remaining h coordinates is $\frac{E(h)}{\binom{n}{h}}$.

Proof: Let $\mathcal{T}$ be the implied $(h, n-h)$ partition, then the required number of codewords is $A^{\mathcal{T}}(h,0)$. The result follows by applying Theorem 3. ■

This result illustrates how the partition weight enumerator of MDS codes is independent of the orientation of the partitions. Since there are $E(h)$ codewords of weight h and there are $\binom{n}{h}$ distinct ways to choose the h zero coordinates, then in such a case one expects that that there are $\frac{E(h)}{\binom{n}{h}}$ codewords for any choice of the h coordinates.

By following the same lines of proof, the result of Theorem 3 can be generalized to an arbitrary number of partitions as in the following theorem:

Theorem 6: For an (n, k, d) *MDS code* $\mathcal{C}$ *with an* $(n_1, n_2, ..., n_p)$ *partition of its coordinates the* p*-partition weight enumerator is given by*

$$A^{\mathcal{T}}(w_1, w_2, ..., w_p) = E(w)\frac{\binom{n_1}{w_1}\binom{n_2}{w_2}....\binom{n_p}{w_p}}{\binom{n}{w}},$$

where $w = \sum_{i=1}^{p} w_i$ *and* $E(w) = |\{\boldsymbol{c} \in \mathcal{C} : \mathcal{W}(\boldsymbol{c}) = w\}|$.

We give numerical examples of PWEs using Theorem 1 and Theorem 6. For these examples, the PWGFs were also verified numerically by generating the $(7, 3, 5)$ RS code.

Example 1: The PWGF for the $(1, 1, 2, 3)$ partition of the coordinates of the $(7, 3, 5)$ RS code over F_8 is

$$\begin{aligned}\mathbb{P}(\mathcal{V}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}) =& 1 + 21\mathcal{V}\mathcal{X}\mathcal{Y}^2\mathcal{Z} + 42\mathcal{V}\mathcal{X}\mathcal{Y}\mathcal{Z}^2 + 21\mathcal{V}\mathcal{Y}^2\mathcal{Z}^2 + 21\mathcal{X}\mathcal{Y}^2\mathcal{Z}^2 + 63\mathcal{V}\mathcal{X}\mathcal{Y}^2\mathcal{Z}^2 \\ &+ 7\mathcal{V}\mathcal{X}\mathcal{Z}^3 + 14\mathcal{V}\mathcal{Y}\mathcal{Z}^3 + 14\mathcal{X}\mathcal{Y}\mathcal{Z}^3 + 42\mathcal{V}\mathcal{X}\mathcal{Y}\mathcal{Z}^3 + 7\mathcal{Y}^2\mathcal{Z}^3 + 21\mathcal{V}\mathcal{Y}^2\mathcal{Z}^3 \\ &+ 21\mathcal{X}\mathcal{Y}^2Z^3 + 217\mathcal{V}\mathcal{X}\mathcal{Y}^2\mathcal{Z}^3.\end{aligned}$$

It could be checked that the sum of the coefficients is the total number of codewords 8^3. For this example, one can also verify the PWGF numerically. □

Example 2: The $(3, 2, 2)$ 3-partition enumerator of the $(7, 5, 3)$ RS code over F_8 is

$$\begin{aligned}\mathbb{P}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}) =& 1 + 7\mathcal{X}^3 + 42\mathcal{X}^2\mathcal{Y} + 70\mathcal{X}^3\mathcal{Y} + 21\mathcal{X}\mathcal{Y}^2 + 105\mathcal{X}^2\mathcal{Y}^2 + 266\mathcal{X}^3\mathcal{Y}^2 \\ &+ 42\mathcal{X}^2\mathcal{Z} + 70\mathcal{X}^3\mathcal{Z} + 84\mathcal{X}\mathcal{Y}\mathcal{Z} + 420\mathcal{X}^2\mathcal{Y}\mathcal{Z} + 1064\mathcal{X}^3\mathcal{Y}\mathcal{Z} + 14\mathcal{Y}^2\mathcal{Z} \\ &+ 210\mathcal{X}\mathcal{Y}^2Z + 1596\mathcal{X}^2\mathcal{Y}^2\mathcal{Z} + 3668\mathcal{X}^3\mathcal{Y}^2\mathcal{Z} + 21\mathcal{X}\mathcal{Z}^2 + 105\mathcal{X}^2\mathcal{Z}^2 \\ &+ 266\mathcal{X}^3\mathcal{Z}^2 + 14\mathcal{Y}\mathcal{Z}^2 + 210\mathcal{X}\mathcal{Y}\mathcal{Z}^2 + 1596\mathcal{X}^2\mathcal{Y}\mathcal{Z}^2 + 3668\mathcal{X}^3\mathcal{Y}\mathcal{Z}^2 \\ &+ 35\mathcal{Y}^2\mathcal{Z}^2 + 798\mathcal{X}\mathcal{Y}^2\mathcal{Z}^2 + 5502\mathcal{X}^2\mathcal{Y}^2\mathcal{Z}^2 + 12873\mathcal{X}^3\mathcal{Y}^2\mathcal{Z}^2.\end{aligned}$$

It can also be verified that $\mathbb{P}(1, 1, 1) = 8^3$. □

Theorem 6 implies that the distribution of the $wE(w)$ non-zero symbols within the codewords of the same Hamming weight w is uniform among the partitions. This issue will be addressed in more detail in the following section.

IV. A RELATIONSHIP BETWEEN COORDINATE WEIGHT AND THE CODEWORD WEIGHT.

In this section, we will show that for MDS codes, one can derive the coordinate weight from the codeword weight. We will discuss whether other linear codes also have this property.

Define $\mathcal{C}_h$ to be the subcode of $\mathcal{C}$ with codewords of Hamming weight h;

$$\mathcal{C}_h \stackrel{\Delta}{=} \{\boldsymbol{c} \in \mathcal{C} : \mathcal{W}(\boldsymbol{c}) = h\}. \tag{27}$$

The following lemma calculates the total weight of any coordinate in the set $\mathcal{C}_h$.

Lemma 7: For an (n, k, d) MDS code $\mathcal{C}$ the total Hamming weight of any coordinate, summed over the subcode $\mathcal{C}_h$, is equal to $\frac{h}{n}E(h)$.

Proof: Let $\mathcal{T}$ be an $(1, n-1)$ partition of $\mathcal{C}$, where the coordinate of choice forms the partition of size one. By Theorem 3, it follows that for any such partition, the number of codewords of $\mathcal{C}$ which are non-zero in this coordinate and have a total weight h, i.e. a weight profile $(1, h-1)$, is

$$A^{\mathcal{T}}(1, h-1) = \frac{\binom{n-1}{h-1}}{\binom{n}{h}} E(h) = \frac{h}{n} E(h). \tag{28}$$

By observing that $A^{\mathcal{T}}(1, h-1)$ is the total weight of the chosen coordinate over codewords in $\mathcal{C}_h$ and that the choice of that coordinate was arbitrary, we are done. ■

This means that the codewords of the subcode $\mathcal{C}_h$, when arranged as the rows of an array, result in a design where the Hamming weight of each row is h and the Hamming weight of each column is $\frac{h}{n}E(h)$. Furthermore, the Hamming distance between any two rows is at least $d = n - k + 1$. We are now ready to prove an important property of MDS codes:

Theorem 8: For an (n, k, d) MDS code $\mathcal{C}$, the ratio of the total weight of any s coordinates of $\mathcal{C}_h$ to the total weight of $\mathcal{C}_h$ is $\frac{s}{n}$. If the s coordinates are 'input' coordinates, then

$$\frac{\sum_{w=1}^{s} w\ O(w, h)}{s} = \frac{h\ E(h)}{n}$$

for any Hamming weight h.

Proof: By Lemma 7, the total weight of any coordinate of $\mathcal{C}_h$ is $(h/n)E(h)$. The total weight of any s coordinates of $\mathcal{C}_h$ is the sum of the weights of the individual coordinates, $s(h/n)E(h)$. By observing that the weight of the s coordinates can be also expressed in terms of the IOWE by $\sum_{w=1}^{s} wO(w, h)$ and $hE(h)$ is the total weight of $\mathcal{C}_h$, the theorem follows. ■

As a side result, we have proven this identity (c.f. (26)):

Corollary 9: Let $\Psi(w)$ be defined as in (25) then

$$\sum_{w} \Psi(w) \binom{s-1}{w-1} \binom{n-s}{h-w} = \Psi(0) \sum_{w} \binom{s-1}{w-1} \binom{n-s}{h-w}.$$

Proof: For an $\mathcal{T} : (s, n-s)$ partition of the coordinates, it follows from Theorem 8 that $\sum_{w=1}^{s} \frac{w}{s} A^{\mathcal{T}}(w, h-w) = \frac{h}{n} E(h) = \binom{n-1}{h-1} \Psi(0)$. Also by Corollary 2, $\sum_{w=1}^{s} \frac{w}{s} A^{\mathcal{T}}(w, h-w) = \sum_{w=1}^{s} \binom{s-1}{w-1} \binom{n-s}{h-w} \Psi(w)$. The proof follows from the identity $\binom{n-1}{h-1} = \sum_{w} \binom{s-1}{w-1} \binom{n-s}{h-w}$. ■

Definition 1: An (n, k) code $\mathcal{C}$ (not necessary MDS) is said to have the multiplicity property $\mathcal{M}$, if for any $\mathcal{T} : (s, n-s)$ partition, $\sum_{w=1}^{s} \frac{w}{s} A^{\mathcal{T}}(w, h-w) = \frac{h}{n} E(h)$ for all Hamming weights h.

We will refer to the partition composed of the s coordinates as the input partition. By Theorem 8, all MDS codes have property $\mathcal{M}$. In general not all linear codes have property $\mathcal{M}$ as seen in the following counter-example:

Example 3: The $(5, 3)$ linear code defined by

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

is composed of the 8 codewords $00000, 10011, 01001, 11010, 00101, 10110, 01100, 11111$. Let the input partition be composed of the first 3 coordinates. For $s = k = 3$, let $\beta(h) = \sum_{w} wO(w, h)$; and $\xi(h) = \frac{3}{5} hE(h)$, then from the following table it is clear that it is not true that this code has property $\mathcal{M}$.

$$\begin{array}{rcccccc} h: & 0 & 1 & 2 & 3 & 4 & 5 \\ \beta(h): & 0 & 0 & 4 & 5 & 0 & 3 \\ \xi(h): & 0 & 0 & 3.6 & 5.4 & 0 & 3 \end{array}.$$

□

It is to be noted that all cyclic codes have property $\mathcal{M}$. This is partially justified by the fact that any cyclic shift of a codeword of weight h is also a codeword of weight h with h/n of the coordinates holding non-zero elements [24]. However, this neither implies Theorem 8 nor is it implied by Theorem 8. For example, an extended RS code is an MDS code but not a cyclic code while an $(7, 4)$ binary Hamming code is cyclic but not MDS. Also, if a code satisfies property $\mathcal{M}$, it is not necessary that the code is either cyclic or MDS. For example, the first order Reed

Muller codes as well as their dual codes, the extended Hamming codes, have property $\mathcal{M}$ but are neither cyclic nor MDS. Next, we discuss some codes with the multiplicity property.

Theorem 10: The first order Reed Muller codes have the multiplicity property $\mathcal{M}$.

Proof: The weight enumerator of the first order Reed Muller codes of length 2^m, $\mathcal{R}(1,m)$, is $\mathbb{E}(\mathcal{W}) = 1 + (2^{m+1}-2)\mathcal{W}^{2^{m-1}} + \mathcal{W}^{2^m}$ and their minimum distance is 2^{m-1}. Let H_{2^m} be the Hadamard matrix of order 2^m and let M be the binary matrix that results from stacking H_{2^m} on top $-\mathrm{H}_{2^m}$ and replacing each $+1$ by 0 and each -1 by 1. (A Hadamard matrix H of order n is an $n \times n$ matrix with entries $+1$ and -1 such that $\mathrm{HH}^T = nI$ and I is the identity matrix. [23, Ch. 18].) The codewords of $\mathcal{R}(1,m)$ are exactly the rows of M [23, Ch. 18]. It follows that each codeword of weight 2^{m-1} has a unique codeword of the same weight which is its binary complement. Thus each coordinate will be equally one and zero in half the number of such codewords. Since the remaining codewords are the all-zero and the all-one codewords, it follows that $\mathcal{R}(1,m)$ has the multiplicity property. ■

We now prove here that if a linear code has property $\mathcal{M}$ then its dual code also has property $\mathcal{M}$. By a straightforward manipulation of the McWilliams identities [3, Ch. 5, Eq. 52] one can show the following relationship between the PWEs of a code and its dual code [25]:

Theorem 11: Let $\mathcal{C}$ be an (n,k) linear code over $\mathbb{F}_q$ and $\mathcal{C}^\perp$ be its dual code. If $\mathcal{T}$ is an $(n1, n2)$ partition of their coordinates, $A(\alpha,\beta)$ and $A^\perp(\alpha,\beta)$ are the PWEs of $\mathcal{C}$ and $\mathcal{C}^\perp$ respectively, then $A(\alpha,\beta)$ and $A^\perp(\alpha,\beta)$ are related by

$$A^\perp(\alpha,\beta) = \frac{1}{|\mathcal{C}|}\sum_{v=0}^{n_2}\sum_{w=0}^{n_1} A(w,v)\mathcal{K}_\alpha(w,n_1)\mathcal{K}_\beta(v,n_2),$$

such that the Krawtchouk polynomial is $\mathcal{K}_\beta(v,\gamma) = \sum_{j=0}^{\beta}\binom{\gamma-v}{\beta-j}\binom{v}{j}(-1)^j(q-1)^{\beta-j}$ for $\beta = 0,1,...,\gamma$.

Define $A_i(\alpha,\beta)$ and $A_i^\perp(\alpha,\beta)$ to be the PWEs for $\mathcal{C}$ and $\mathcal{C}^\perp$ respectively when an $(1, n-1)$ partition is applied to their coordinates such that the first partition of cardinality one is composed of the ith coordinate.

Theorem 12: An (n,k) linear code over $\mathbb{F}_q$ has the multiplicity property iff its dual code has the multiplicity property.

Proof: Let $\mathcal{C}$ be an (n,k) linear code over $\mathbb{F}_q$ with property $\mathcal{M}$ and an $(1,n-1)$ PWE

$A_i(\alpha, \beta)$. From Theorem 11 the PWE of the dual code $\mathcal{C}^\perp$ is

$$A_i^\perp(1,\beta) = \frac{1}{|\mathcal{C}|}\sum_{v=0}^{n-1}\sum_{w=0}^{1} A_i(w,v)\mathcal{K}_1(w,1)\mathcal{K}_\beta(v,n-1). \tag{29}$$

Since $\mathcal{C}$ has property $\mathcal{M}$, then $A_i(1,v) = \frac{v+1}{n}E_{\mathcal{C}}(v+1)$ and $A_i(0,v) = E_{\mathcal{C}}(v) - A_i(1,v-1) = (1-\frac{v}{n})E_{\mathcal{C}}(v)$. By substituting in (29), it follows that $A_i^\perp(1,\beta) = A_j^\perp(1,\beta)$ for any $i,j \in \{1,2,...,n\}$ and $\sum_{i=1}^n A_i^\perp(1,\beta) = nA_i^\perp(1,\beta)$ for any i. Counting the total weight of the codewords in $\mathcal{C}^\perp$ with Hamming weight h by two different ways, we get $\sum_{i=1}^n A_i^\perp(1,\beta) = (\beta+1)E_{\mathcal{C}^\perp}(\beta+1)$. It follows that $A_i^\perp(1,\beta) = \frac{\beta+1}{n}E_{\mathcal{C}^\perp}(\beta+1)$ and $\mathcal{C}^\perp$ has property $\mathcal{M}$.

For the converse, assume that $\mathcal{C}$ does not satisfy property $\mathcal{M}$ but $\mathcal{C}^\perp$ does. From the previous argument $(\mathcal{C}^\perp)^\perp$ has property $\mathcal{M}$. Since for linear codes $(\mathcal{C}^\perp)^\perp = \mathcal{C}$, we reach a contradiction. ■

Since the dual codes of MDS codes are also MDS codes, this result strengthens Theorem 8. This theorem somehow strengthens the result of Theorem 8 since the dual codes of MDS codes are again MDS codes. The dual codes of cyclic codes are also cyclic codes. One can also use this theorem to show that certain codes have the multiplicity property.

Corollary 13: The extended Hamming codes have property $\mathcal{M}$.

Proof: An extended Hamming code of length 2^m is the dual of the first order RM code $\mathcal{R}(1,m)$ [3], which by Theorem 10 has property $\mathcal{M}$. ■

Extended Hamming codes also have transitive automorphism groups [26] which gives another proof to Corollary 13. Some product codes also have the multiplicity property [26], [27].

V. Average binary image of Reed Solomon codes

The binary image $\mathcal{C}^b$ of an (n,k) code $\mathcal{C}$ over F_{2^m} is obtained by representing each symbol by an m-dimensional binary vector in terms of a basis of the field [22]. The weight enumerator of $\mathcal{C}^b$ will vary according to the basis used. In general, it is also hard to know the weight enumerator of the binary image of a certain Reed Solomon code obtained by a specific basis representation (e.g. [15], [16]). For performance analysis, one could average the performance over all possible binary representations of $\mathcal{C}$. By assuming that the all such representations are equally probable, it follows that the distribution of the bits in a non-zero symbol follows a binomial distribution and the probability of having i ones in a non-zero symbol is $\frac{1}{2^m-1}\binom{m}{i}$. The generating function

of the *average* weight enumerator of the binary image of a non-zero symbol is

$$F(\mathcal{Z}) = \sum_{i=1}^{m} \frac{1}{2^m - 1} \binom{m}{i} \mathcal{Z}^i = \frac{(1+\mathcal{Z})^m - 1}{2^m - 1}, \tag{30}$$

where the power of x denotes the binary weight and the all zero vector is excluded since the binary weight of a non-zero symbol is at least one. Suppose a codeword has w non-zero symbols, and the distribution of the ones and zeros in each symbol is independent from other symbols, then the possible binary weight, b, of this codeword ranges from w to mw. Since there are $E(w)$ codewords with symbol Hamming weight w, then the *average binary* weight generating function can be derived by

$$\begin{aligned} \tilde{\mathbb{E}}_{\mathcal{C}^b}(\mathcal{X}) &= \sum_{b=0}^{nm} \tilde{E}(b) \mathcal{X}^b & (31)\\ &= \mathbb{E}_{\mathcal{C}}(\mathcal{X}) \big|_{\mathcal{X}:=F(\mathcal{X})} & (32)\\ &= \sum_{h=0}^{n} \frac{E(h)}{(2^m-1)^h} \left((1+\mathcal{X})^m - 1\right)^h . & (33) \end{aligned}$$

A closed form formula for the average binary weight enumerator (BWE) is

$$\begin{aligned} \tilde{E}(b) &= \text{Coeff}\left(\tilde{\mathbb{E}}_{\mathcal{C}^b}(\mathcal{X}), \mathcal{X}^b\right) & (34)\\ &= \sum_{w=d}^{n} \frac{E(w)}{(2^m-1)^w} \sum_{j=0}^{w} (-1)^{w-j} \binom{w}{j} \binom{jm}{b}; \quad b \geq d. & (35) \end{aligned}$$

These results apply to any maximum distance separable code defined over $\mathbb{F}_q$, where $q = 2^m$ and not necessarily an RS code. Widely used RS (MDS) codes have a code length $n = 2^m - 1$. In that case the BWE derived in (34) agrees with the average BWE of a class of GRS codes [17]. In other words two ensembles have the same weight enumerator; the first ensemble is the ensemble of all possible binary images of a specific RS code, the second ensemble is the binary image (with a specific basis representation) of the ensemble of generalized RS codes derived from the original RS code by multiplying each column in the generator matrix by some non-zero element in the field. It is easy to see that $G_o = 1$ and that $\tilde{E}(b) = 0$ for $0 < b < d$. By substituting for $E(w)$, for $b \geq d$, the binary weight enumerator (BWE) is given by

$$\tilde{E}(b) = (q-1) \sum_{w=d}^{n} \left(\frac{q}{q-1}\right)^w \binom{n}{w} \sum_{v=0}^{w-d} (-1)^v \binom{w-1}{v} \left[\sum_{j=\lceil b/m \rceil}^{w} (-1)^{w-j} \binom{w}{j} \binom{jm}{b} q^{-(d+v)} \right]. \tag{36}$$

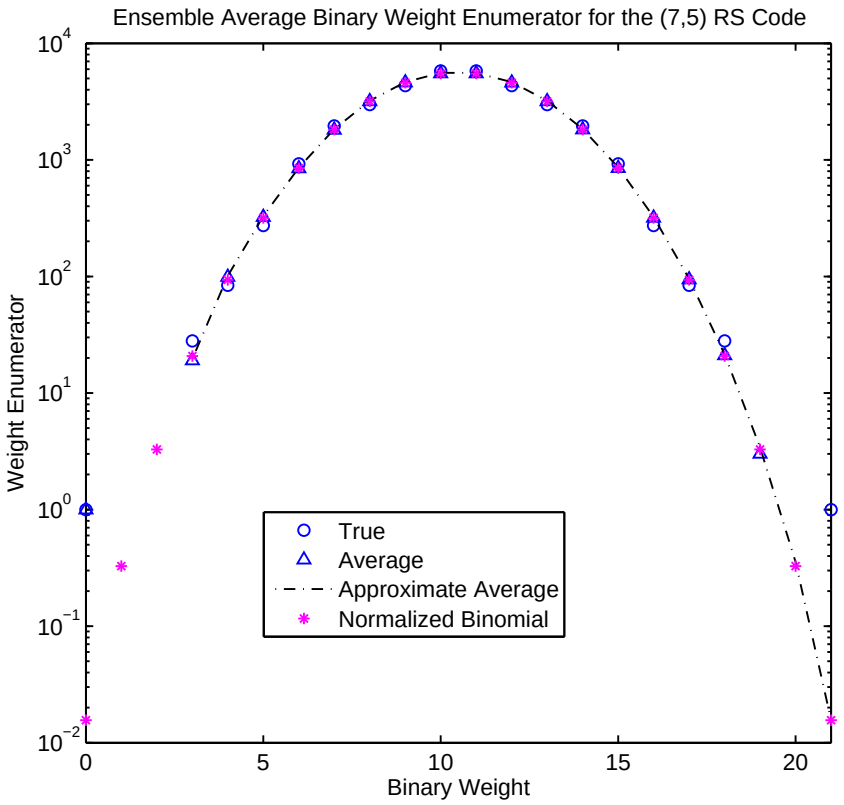

Fig. 4. True BWE versus the averaged BWE for the (7,5) RS code over $\mathbb{F}_8$

Although it is easy to evaluate the above formula, the term $\binom{jm}{b}$ may diverge numerically for large j. Using the Stirling approximation for $\binom{jm}{b}$ [3], $\tilde{E}(b)$ could be approximated as

$$\tilde{E}(b) \approx \sum_{w=d}^{n}(q-1)\left(\frac{q}{q-1}\right)^{w}\binom{n}{w}\sum_{v=0}^{w-d}(-1)^{v}\binom{w-1}{v}\sum_{j=\lceil b/m\rceil}^{w}\mathcal{F}(j), \tag{37}$$

where

$$\mathcal{F}(j)=\begin{cases}(-1)^{w-j}\binom{w}{j}2^{\lambda(j)}; & j > b/m\\ \\ (-1)^{w-j}\binom{w}{j}2^{-m(d+v)}; & j = b/m\end{cases}, \tag{38}$$

and $\lambda(j) = m(jH(\psi_{b,j}) - d - v) - \frac{1}{2}\log_2\left(2\pi jm\psi_{b,j}(1-\psi_{b,j})\right)$ for $\psi_{b,j} = b/jm$ and $q = 2^m$. These bounds could be further simplified (and thus loosened) by observing that for $n \leq q-1$,

$$1 \leq \left(\frac{q}{q-1}\right)^{w} \leq \left(\frac{q}{q-1}\right)^{q-1} \leq \lim_{q\to\infty}\left(\frac{q}{q-1}\right)^{q-1} = \mathbf{e} \tag{39}$$

and substituting in (37).

In Fig. 4, the averaged BWE and the true BWE for a specific basis representation found by computer search are plotted for the $(7,5)$ RS code over $\mathbb{F}_8$. The average weight enumerator of (36) is labeled "Average" while the approximation of (37) is labeled 'Approximate Average'. It is observed that a good approximation of the average binary weight enumerator for $h \geq d$ is the normalized binomial distribution which corresponds to a random code with the same dimension over $\mathbb{F}_q$

$$\tilde{E}(h) \approx q^{-(n-k)} \binom{mn}{h}. \tag{40}$$

This observation can be somehow justified by the central limit theorem, where the binary weight of a codeword is a random variable which is the sum of n independent random variables corresponding to the binary weights of the symbols. For large n, the distribution of the binary weight is expected to converge to that of random codes. The following theorem shows that the average BWE can be upper bounded by a $\left(\frac{q}{q-1}\right)^{(n-k)}$ multiple of the above approximation.

Theorem 14: The average binary weight enumerator is upper bounded by

$$\tilde{E}(h) \leq (q-1)^{-(n-k)} \binom{mn}{h}.$$

Proof: An upper bound on the symbol weight enumerator of an (n,k,d) MDS code defined over $\mathbb{F}_q$ is [28, Eq. 12]

$$E(w) \leq \binom{n}{w}(q-1)^{w-d+1}; \quad w \geq d. \tag{41}$$

Substituting in (34) it follows that for $b \geq d$

$$\tilde{E}(b) \leq (q-1)^{k-n} \sum_{w=d}^{n} \binom{n}{w} \left[\sum_{j=\lceil b/m \rceil}^{w} (-1)^{w-j} \binom{w}{j} \binom{jm}{b} \right]. \tag{42}$$

By doing a change of variables $\alpha = mj$ and changing the order of summations

$$\begin{aligned}
\tilde{E}(b) &\leq (q-1)^{k-n} \sum_{w=d}^{n} \sum_{\alpha=b}^{mw} (-1)^{w-j} \binom{n}{w} \binom{w}{\alpha/m} \binom{\alpha}{b} \\
&= (q-1)^{k-n} \sum_{\alpha=b}^{nm} (-1)^{-\frac{\alpha}{m}} \binom{\alpha}{b} \sum_{w=\max(\frac{\alpha}{m},d)}^{n} (-1)^{w} \binom{n}{w} \binom{w}{\alpha/m} \\
&\leq (q-1)^{k-n} \sum_{\alpha=b}^{nm} (-1)^{-\frac{\alpha}{m}} \binom{\alpha}{b} \sum_{w=\frac{\alpha}{m}}^{n} (-1)^{w} \binom{n}{w} \binom{w}{\alpha/m}.
\end{aligned}$$

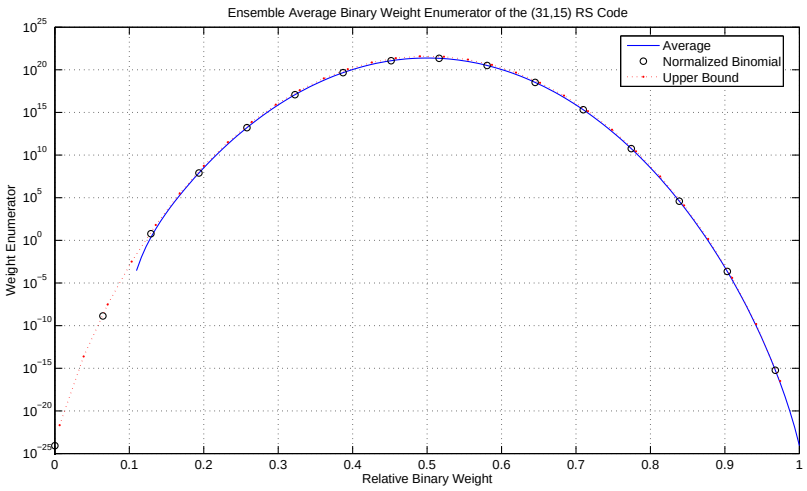

Fig. 5. For the $(31, 15)$ RS code over $\mathbb{F}_{32}$, the ensemble average weight enumerator of (34) is compared with the random code ensemble (40) and the upper bound of Theorem 14. They are labeled 'Average', 'Normalized Binomial' and 'Upper Bound' respectively.

From the identity $\binom{n}{m}\binom{m}{p} = \binom{n}{p}\binom{n-p}{m-p}$ it follows that $\sum_{k=m}^{n}(-1)^k\binom{n}{k}\binom{k}{m} = (-1)^m\delta_{nm}$ where $\delta_{n,m}$ is the Kronecker delta function. It follows that

$$\begin{aligned}\tilde{E}(b) &\leq (q-1)^{k-n}\sum_{\alpha=b}^{nm}\binom{\alpha}{b}\delta_{\frac{\alpha}{m},n}\\ &= (q-1)^{k-n}\binom{mn}{b},\end{aligned}$$

which completes the proof. ■

In Fig. 5, we plot the ensemble average weight enumerator of (34) and compare it with the weight enumerator of a random code with the same dimension (40). We also compare it with the simple upper bound of Theorem 14. It is observed that the upper bound of Theorem 14 is fairly tight and that a good approximation for the ensemble weight enumerator is that of random codes. In fact, as length of the code (and the size of the finite field) tend to infinity

$$\tilde{E}(h) \leq \left(\frac{q}{q-1}\right)^{(n-k)} q^{-(n-k)}\binom{mn}{h} \tag{43}$$

$$\leq \mathbf{e}2^{-m(n-k)}\binom{mn}{h} \tag{44}$$

$$\leq \frac{\mathbf{e}}{\sqrt{2\pi mn\lambda(1-\lambda)}}\,2^{mn(H_2(\lambda)-1+R)}, \tag{45}$$

where $b = \lambda mn$, $R = k/n$ is the code rate and $H_2(\lambda)$ is the binary entropy function. The last inequality follows from the Stirling's inequality [3, p. 309]. Let the asymptotic weight enumerator exponent of a code $\mathcal{C}$, of length N and weight enumerator $E_{\mathcal{C}}$, to be defined as

$$\Xi(\lambda) \triangleq \lim_{N\to\infty} \frac{\log_2\left(E_{\mathcal{C}}(\lambda N)\right)}{N}. \tag{46}$$

It follows that the asymptotic weight enumerator exponent of the ensemble of binary images of Reed-Solomon codes is

$$\begin{aligned}
\tilde{\Xi}(\lambda) &= \lim_{\substack{n\to\infty \\ m\to\infty}} \frac{\log_2\left(\tilde{E}(\lambda mn)\right)}{mn} \\
&\leq \lim_{\substack{n\to\infty \\ m\to\infty}} \frac{\log_2(\mathbf{e}) - \frac{1}{2}\log_2(mn) - \frac{1}{2}\log_2(2\pi\lambda(1-\lambda))}{mn} + H_2(\lambda) - 1 + R \\
&= H_2(\lambda) - (1-R).
\end{aligned} \tag{47}$$

In other words, as the code length and the finite field size tend to infinity, the weight enumerator of the ensemble of binary images of an RS code approaches that of a random code.

The error correcting capability of a code relies a lot on the minimum distance of the code, which will be analyzed in the next section.

VI. The Binary Minimum Distance of the Ensemble of Binary Images of Reed-Solomon Codes

The error correcting capability of a code relies a lot on the minimum distance of the code. We will now consider the minimum distance of the ensemble of binary images of a certain Reed Solomon code. The average minimum distance of the binary image of the RS code could be defined to be the smallest weight b whose average BWE $\tilde{E}(b)$ is greater than or equal to one (note that $\tilde{E}(b)$ is a real number). Let d_b be the average BMD, then

$$d_b \triangleq \inf_{b\geq d}\{b : \tilde{E}(b) \geq 1\}. \tag{48}$$

The number d_b could be found exactly by numerical search. However, it will also be useful to find a lower bound on d_b. It is straight forward to note that the binary minimum distance (BMD) is at least as large as the symbol minimum distance d;

$$d_b \geq n - k + 1. \tag{49}$$

In the following theorems, we will give some lower bounds on the average binary minimum distance of the ensemble of binary images.

Theorem 15: The minimum distance of the ensemble of binary images of an (n, k, d) *RS code over* $\mathbb{F}_{2^m}$ *is lower bounded by*

$$d_b \geq \inf_{b \geq d} \left\{ b : \binom{mn}{b} \geq (2^m - 1)^{n-k} \right\}.$$

Proof: From the upper bound on $\tilde{E}_b$ of Theorem 14, and the definition of d_b, the theorem follows. ■

By taking only the term corresponding to $j = w$ in the alternating sign summation in (42), one can show that an upper bobound on the minimum distance of Theorem 16 will not be tighter than that of Theorem 15. und on the ensemble weight enumerator is

$$\tilde{E}(b) \leq (q-1)^{k-n} \sum_{w=d}^{n} \binom{n}{w} \binom{wm}{b} \tag{50}$$

Theorem 16: A lower bound on d_b *is*

$$d_b \geq \inf_{b \geq d} \left\{ b : \sum_{w=d}^{n} \binom{n}{w} \binom{wm}{b} \geq (2^m - 1)^{n-k} \right\}.$$

Proof: By taking only the term corresponding to $j = w$ in the alternating sign summation in (42), it follows that

$$\tilde{E}(b) \quad \leq \quad (q-1)^{k-n} \sum_{w=d}^{n} \binom{n}{w} \binom{wm}{b}.$$

The theorem follows from the definition of d_b. ■

Since the upper bound on the weight enumerator of (50) is not tighter than the bound of Theorem 14, it is expected that the lower bound on the minimum distance of Theorem 16 will not be tighter than that of Theorem 15.

Since the binary minimum distance of the ensemble is at least as large as the symbol minimum distance (cf. 49), it is interesting to determine when the binary minimum distance is equal to the symbol minimum distance which is linear in the rate R of the code.

Lemma 17: The average binary minimum distance of an MDS code over $\mathbb{F}_{2^m}$ *is equal to its symbol minimum distance for all rates greater than or equal to* $R_o = 1 - \frac{d_o - 1}{n}$ *where* d_o *is the largest integer* d' *such that*

$$\frac{1}{d'} \log_2 \left((2^m - 1) \binom{n}{d'} \right) \geq \log_2(2^m - 1) - \log_2(m). \tag{51}$$

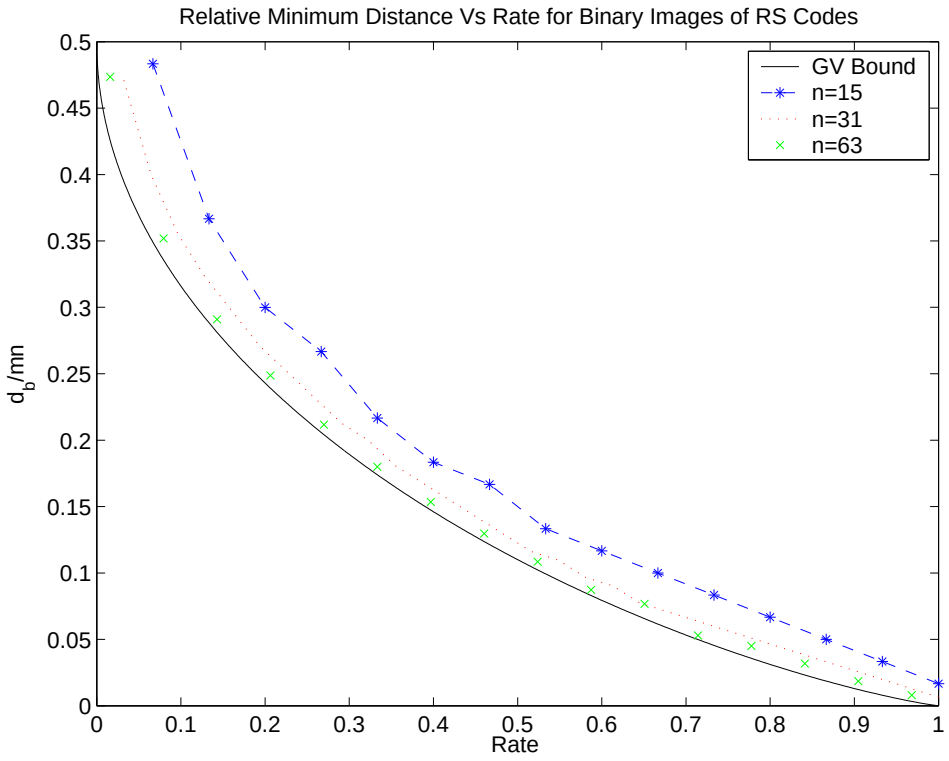

Fig. 6. Relative binary minimum distance for the ensemble of binary images of Reed Solomon codes, of lengths 15, 31 and 63 over finite fields of sizes 16, 32 and 64 respectively, plotted versus the code rate and compared with the Gilbert-Varshamov bound.

Proof: The number of codewords in an MDS code with symbol weight $d = n - k + 1$ is $E(d) = (q-1)\binom{n}{d}$. The binary image could be of binary weight d only if the codeword is of symbol weight d and the binary representation of each non-zero symbol has only one non-zero bit. This happens with probability $\left(\frac{m}{2^m-1}\right)^d$, where $m = \log_2(q)$. So the average number of codewords with binary weight d is

$$\tilde{E}(d) = E(d)\left(\frac{m}{2^m-1}\right)^d = (q-1)\binom{n}{d}\left(\frac{\log_2(q)}{q-1}\right)^d. \tag{52}$$

From the definition of the average binary minimum distance, the lemma follows. ■

Asymptotically, it could be shown that R_o is the smallest rate such that

$$\frac{H_2(1-R_o)}{(1-R_o)} \geq \log_2(n) - \log_2(\log_2(n)), \tag{53}$$

where $n \approx q$ and

$$H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x) \tag{54}$$

is the binary entropy function. This implies that the rate R_o, at which the symbol minimum distance is equal to the ensemble binary minimum distance, tends to one as the length of the code tends to infinity.

The Gilbert-Varshamov (GV) bound is defined by [3],

$$\lim_{n\to\infty} \{R(\delta) - (1 - H_2(\delta))\} \geq 0 \ \text{ for } \ 0 < \delta < \frac{1}{2}, \tag{55}$$

where $\delta = d_b/(mn)$ is the ratio of the binary minimum distance to the total length of the code and $R(\delta)$ is rate of the code with a relative minimum distance δ. Retter showed that for sufficiently large code lengths, most of the codes in the binary image of the ensemble of generalized RS codes lie close to the GV bound by showing that the number of codewords with weights lying below the GV bound in all generalized RS codes of the same length and rate are less than half the number of such generalized RS codes [17]. Next, we show a related result for the ensemble of binary images of an RS code, with a binary weight enumerator $\tilde{E}(b)$.

We will now determine a bound on the asymptotic relative binary minimum distance (as the length tends to infinity) of the ensemble of binary images, δ_∞

$$\delta_\infty \stackrel{\Delta}{=} \inf_{\lambda} \{\tilde{\Xi}(\lambda) \geq 0\}. \tag{56}$$

From the asymptotic analysis of (47), we showed that

$$\tilde{\Xi}(\lambda) \leq H_2(\lambda) - (1 - R). \tag{57}$$

It thus follows that

$$\delta_\infty \geq \inf_{\lambda} \{H_2(\lambda) \geq (1 - R)\}. \tag{58}$$

One can then deduce that

$$H_2(\delta_\infty) - (1 - R(\delta_\infty)) \geq 0. \tag{59}$$

In other words, we have proved the following theorem,

Theorem 18: The ensemble of binary images of an Reed Solomon code asymptotically satisfies the Gilbert Varshamov bound.

This is not very surprising since we have shown that the ensemble average behaves like a binary random code. Note that this is for the average binary image of the RS code and not for a specific valid binary image. Since this theorem is for the ensemble average, it might imply that

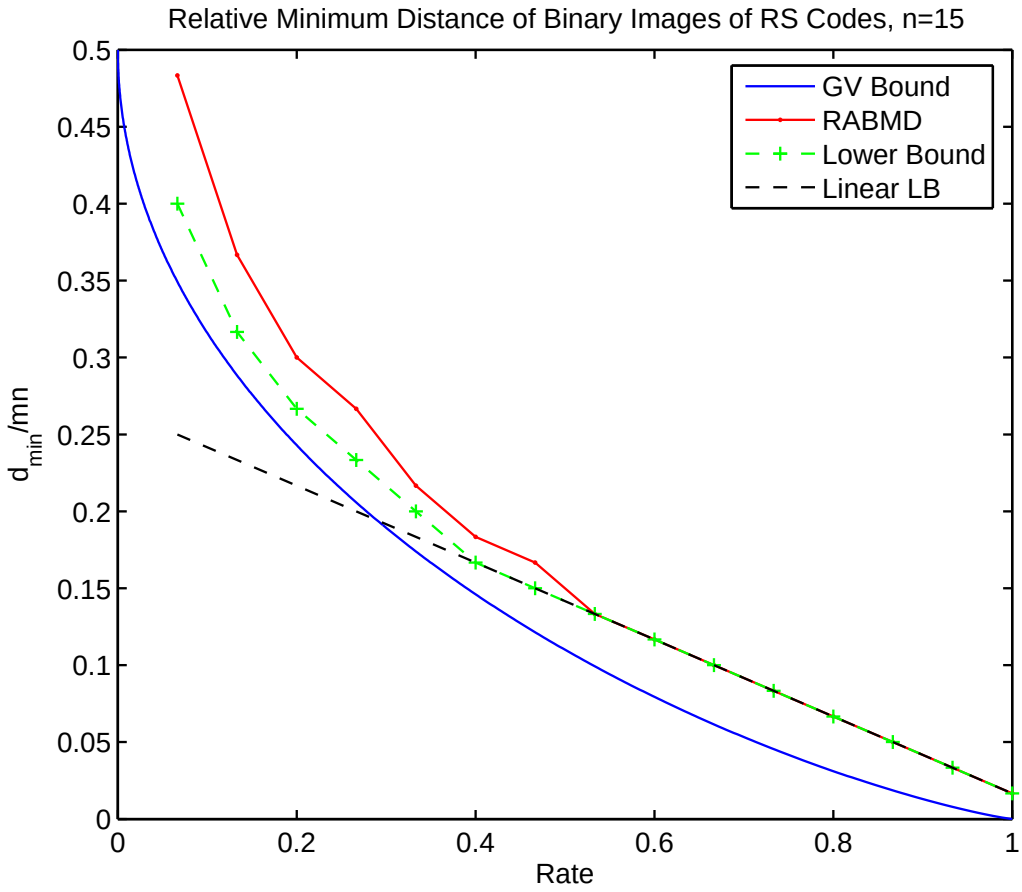

Fig. 7. The relative binary minimum distance for the ensemble of binary images of Reed Solomon codes, of length 15 over $\mathbb{F}_{16}$ plotted versus the code rate. The numerical minimum distance (48) is labeled 'RABMD' and compared with the lower bounds of Theorem 15 and (49) which are labeled 'Lower Bound' and 'Linear LB' respectively. The Gilbert-Varshamov bound is plotted and labeled 'GV Bound'.

some codes in the ensemble may have a minimum distance asymptotically satisfying the GV bound. However, we do not know of a specific code in the ensemble that satisfies the bound.

In Fig. 6, we show the relative average binary minimum distance for binary images of Reed Solomon codes, calculated numerically by (48), for different code lengths. It is observed that as the length and the size of the finite field increases, the relative minimum distance decreases. From Theorem 18, the relative binary minimum distance should approach the GV bound as the length tends to infinity. In Fig. 7 and Fig. 8, we study the relative average binary minimum distance for code lengths $n = 15$ and $n = 31$ respectively. We compare it with the Gilbert-Varshamov

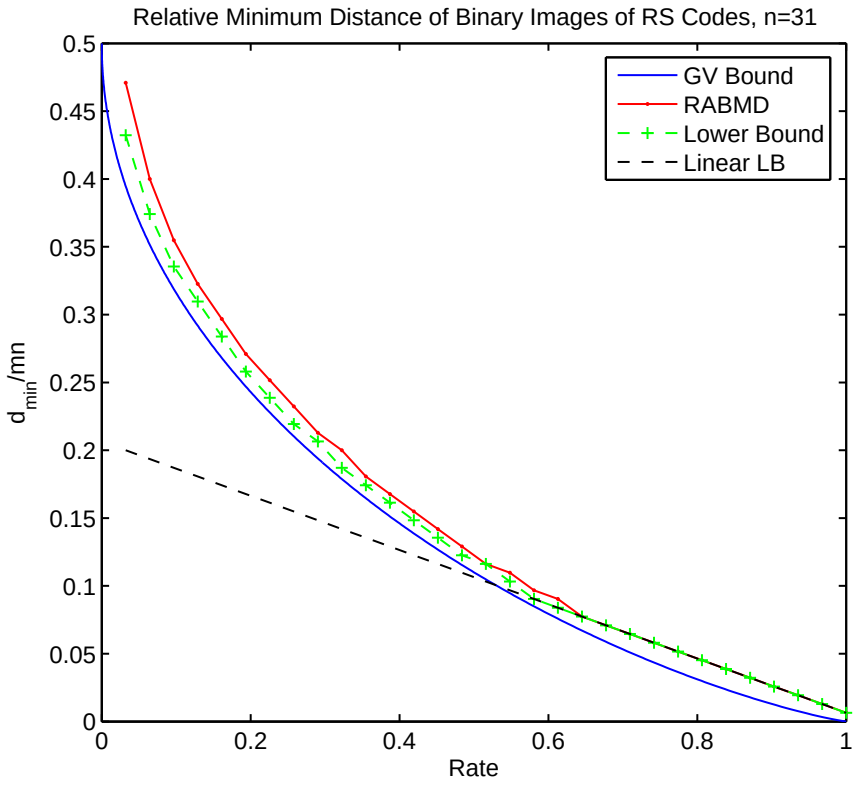

Fig. 8. The relative binary minimum distance for the ensemble of binary images of Reed Solomon codes, of length 31 over $\mathbb{F}_{32}$ plotted versus the code rate. The numerical minimum distance (48) is labeled 'RABMD' and compared with the lower bounds of Theorem 15 and (49) which are labeled 'Lower Bound' and 'Linear LB' respectively. The Gilbert-Varshamov bound is plotted and labeled 'GV Bound'.

bound and the lower bounds of Theorem 15 and the linear bound of (49). We observe that the lower bound of Theorem 15 is pretty tight and it provides a simple way to evaluate the minimum distance of the ensemble. Moreover it is always lower bounded by the GV bound. By comparing with the linear lower bound of (49), it is noticed that for $n = 15$ and $k \geq 8$, the average BMD is equal to the symbol minimum distance, d, as expected from Lemma 17. As the rate decreases, this linear lower bound becomes very loose and the average binary minimum distance exceeds the symbol minimum distance.

VII. Performance of the Maximum Likelihood Decoders

Let $\boldsymbol{c}$ be the binary image of a codeword in the (n, k, d) RS code $\mathcal{C}$. The binary phase shift keying (BPSK) modulated image of $\boldsymbol{c}$ is $\boldsymbol{x} = \mathcal{M}(\boldsymbol{c}) = 1 - 2\boldsymbol{c}$. This will be transmitted over a standard binary input additive white Gaussian noise (AWGN) channel. The received vector is $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}$, where $\boldsymbol{z}$ is an AWGN vector. Since the considered codes are linear, it is safe to assume that the all zero codeword (in fact its binary image) is transmitted. Hard-decision is done to the received bits to obtain the vector $\bar{\boldsymbol{y}}$ where $\bar{y}_i = \frac{1-\text{sign}(y_i)}{2}$ and the HD-ML decoder's output is the codeword $\hat{\boldsymbol{c}}$ such that

$$\hat{\boldsymbol{c}} = \arg \min_{\boldsymbol{v} \in \mathcal{C}^b} d(\bar{\boldsymbol{y}}, \boldsymbol{v}) \tag{60}$$

where $d(\boldsymbol{u}, \boldsymbol{v})$ is the (binary) Hamming distance between $\boldsymbol{u}$ and $\boldsymbol{v}$. This is equivalent to transmitting the codeword $\boldsymbol{c}$ through a binary symmetric channel (BSC) with cross over probability $p = Q(\sqrt{2R\gamma})$ where γ is the bit signal to noise ratio and R is the code rate.

As discussed before, bounds on the error probability of linear codes require the knowledge of the weight enumerator. For a specific binary image, it is very hard to know the weight enumerator. It is also hard to agree on the use of a specific binary image or to speculate which binary image has been used. So the question we really need to answer is the expected performance if any binary image of a specific RS code is used. Our approach is to consider the binary code of a weight enumerator equal to the ensemble average weight enumerator.

The performance of the hard-decision maximum likelihood (HD-ML) decoder can be upper bounded with the well known union bound by resorting to the average weight enumerator of the ensemble

$$P(\mathcal{E}_{HML}) \leq \sum_{b=d_b}^{mn} \tilde{E}(b) \sum_{w=\lceil \frac{b}{2} \rceil}^{b} \binom{b}{w} p^w (1-p)^{b-w}, \tag{61}$$

where $P(\mathcal{E}_{HML})$ denotes the codeword error probability of the HD-ML decoder. Alternatively, one could use the ensemble average weight enumerator with tighter bounds. The best well known upper bound on the performance of a HD-ML decoding of linear codes on binary symmetric channels is the Poltyrev bound [29].

The soft-decision maximum likelihood decoder solves the following optimization problem,

$$\hat{\boldsymbol{c}} = \arg \min_{\boldsymbol{v} \in \mathcal{C}_b} \|\boldsymbol{y} - \mathcal{M}(\boldsymbol{v})\|^2 \tag{62}$$

where $\|\boldsymbol{x}\|$ is the Euclidean norm of $\boldsymbol{x}$. Assuming that the all-zero codeword is BPSK modulated and transmitted over a memoryless AWGN channel, the probability that a certain codeword of binary weight b is chosen at the decoder instead of the transmitted all-zero codeword is [30, Eq. 8.1-49] $P_b = Q\left(\sqrt{2\gamma R b}\right)$, where γ is the signal to noise ratio (SNR) per bit and $R = k/n$.

Then a heuristic union lower bound on the codeword error probability of the soft-decision maximum-likelihood decoder (specifically true at high SNRs) is the probability that a codeword of minimum weight d_b is erroneously decoded,

$$P(\mathcal{E}_{SML}) \gtrsim \tilde{E}(d_b) Q\left(\sqrt{2\gamma R d_b}\right). \tag{63}$$

A union upper bound on the codeword error probability is the sum of all possible errors,

$$P(\mathcal{E}_{SML}) \leq \sum_{b \geq d_b} \tilde{E}(b) Q\left(\sqrt{2\gamma R b}\right). \tag{64}$$

The union bound is loose at low SNRs. Poltyrev described a tangential sphere bound (TSB) on the error probability of binary block codes BPSK modulated in AWGN channels [29]. This is a very tight upper bound on the ML error probability. We use it in conjunction with the average binary weight enumerator to find a tight upper bound on the error probability of ML decoding of RS codes. Divsalar also introduced in [31] a simple tight bound (that involves no integrations) on the error probability of binary block codes, as well as a comparison of other existing bounds.

The Berlekamp-Massey (BM) decoder is a symbol-based hard-decision decoder which can correct a number of symbol errors upto half the minimum distance of the code, $\tau_{BM} = \lfloor \frac{n-k}{2} \rfloor$. The error plus failure probability of the BM decoder has been well studied [28], [32] and can be simply given by

$$P(\mathcal{E}_{BM}) = 1 - \sum_{j=0}^{\tau_{BM}} \binom{n}{j} (1-s)^j s^{n-j},$$

where s is the probability that a symbol is correctly received $s = \left(1 - Q\left(\sqrt{2\gamma R}\right)\right)^m$. The Guruswami-Sudan decoder is also a symbol-based HD decoder but can correct more than half the minimum distance of the code $\tau_{GS} = \lceil n - \sqrt{nk} - 1 \rceil$. The performance of a hard-decision 'sphere' decoder that corrects any number of $\tau \geq \tau_{BM}$ symbol errors as well that of the corresponding maximum likelihood decoder over q-ary symmetric channels have been recently analyzed [33], [34].

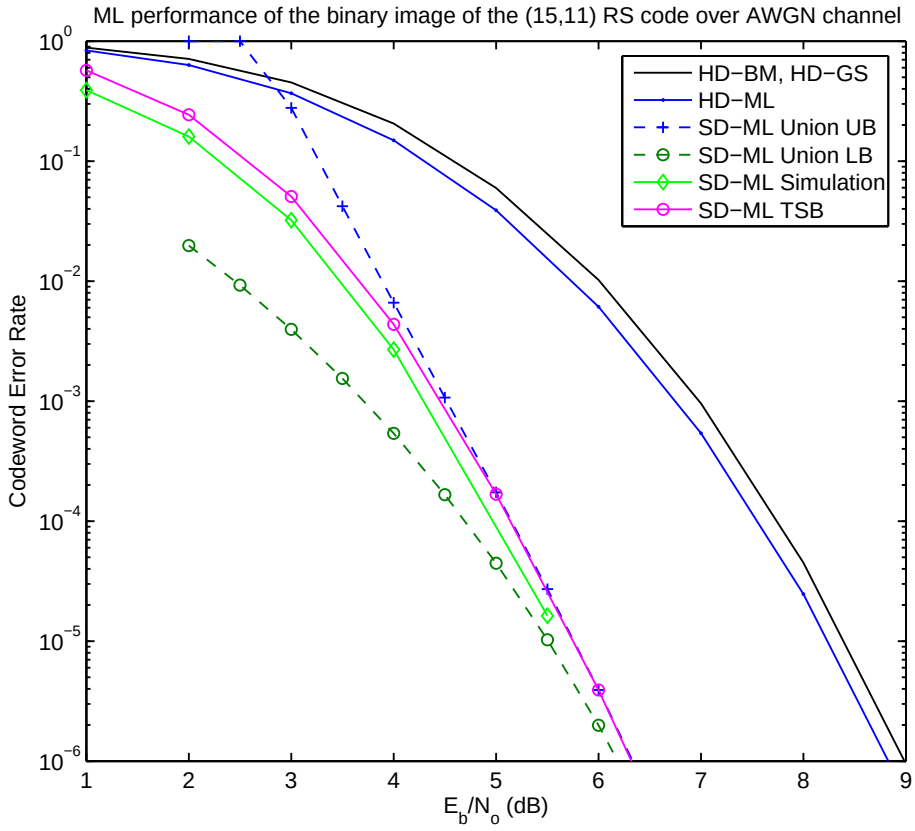

Fig. 9. Performance bounds of the binary image of the $(15, 11)$ RS code over $\mathbb{F}_{16}$ when transmitted over a binary input AWGN channel: The analytic performance of the symbol-level hard-decision Berlekamp-Massey and Guruswami-Sudan decoders are shown and are labeled by 'HD-BM' and 'HD-GS' respectively. These are in turn compared to the bit-level HD ML decoder labeled 'HD-ML'. The union upper bound (64), lower bound (63) and the tangential sphere bound on the soft-decision ML error probability are labeled 'SD-ML Union UB', 'SD-ML Union LB' and 'SD-ML TSB' respectively. The simulated performance of an SD ML decoder is labeled 'SD-ML Simulation'.

We evaluate the average performance of RS codes when its binary image is BPSK modulated and transmitted over an AWGN channel. In Fig. 9, we consider a specific binary image of the $(15, 11)$ RS code over $\mathbb{F}_{16}$. Soft-decision maximum likelihood decoding was simulated using the BCJR algorithm [35] on the trellis associated with the binary image of the RS code [36]. By comparing this with the average TSB, we observe that our technique for bounding the performance of the soft-decision ML decoder provides tight upper bounds on the actual performance of a specific binary image. It is clear that at low SNRs the (averaged) TSB give a close approximation of the ML error probability. By comparing this bound with the union

upper and lower bounds of (64) and (63), we observe that the TSB coincides with the union bounds at high SNRs. As from (63), the union lower bound is characterized by the minimum distance term. Indeed, the SNR at which the performance of the maximum likelihood decoder is dominated by the minimum distance term was recently studied by Fossorier and was termed the *critical point* for ML decoding [37]. The decoding radius of the GS decoder is the same as that of the BM decoder for the $(15, 11)$ code, which is of relatively high rate. However, their performance is very close to that of the HD-ML decoder.

In Fig. 10, we consider the performance of the binary image of the $(31, 15)$ RS code over $\mathbb{F}_{16}$ when BPSK modulated and transmitted over an AWGN channel. We compare the performance of a bit-level HD-ML decoder with that of a symbol-level HD-ML decoder by deploying the bounds of [29] and [34] respectively. The symbol-level decoder operates by first grouping m bits to symbols in $\mathbb{F}_{2^m}$ after hard-decision. It seems that for this half-rate code, the performance of a bit-level HD decoder is better than the corresponding symbol-level decoder (about 1.5 dB coding gain). We also compare the performance with that of the symbol-level HD-BM and the HD-GS algorithms. For the $(31, 15)$ code, bit-level HD-ML decoding has more than 2 dB gain over the BM decoder, whereas SD-ML decoding offers another 2 dB gain over bit-level HD-ML decoding. The SD-ML decoder has about 4 dB gain over the BM decoder and 2 dB gain over the HD-ML decoder. Bounds on the performance of the maximum likelihood decoder provides a benchmark to compare the performance of other suboptimum algorithms. To emphasize this, the performance of a bit-level soft-decision decoder, developed by El-Khamy and McEliece [12], acting on a specific binary image is also plotted. Only by comparing it to the SD-ML bound can one conclude that this soft-decision algorithm operates within 1 dB of the optimum soft-decision algorithm.

VIII. Binary Partition Weight Enumerator of MDS Codes

In this section, we study the partition weight enumerator of the binary image of an RS (MDS) code. Let $\mathcal{T}$ be a partition of the coordinates of an MDS code $\mathcal{C}$ defined over $\mathbb{F}_{2^m}$. Let $\mathcal{T}_b$ be the partition of the coordinates of the code's binary image $\mathcal{C}^b$ implied by $\mathcal{T}$ when each symbol is represented with its binary image. The number of the partitions in $\mathcal{T}$ and $\mathcal{T}_b$ is the same but the size of each partition is m times larger. This is illustrated by example in Figure 11. The *binary partition weight enumerator* (PWE) gives the number of codewords in the binary image with a

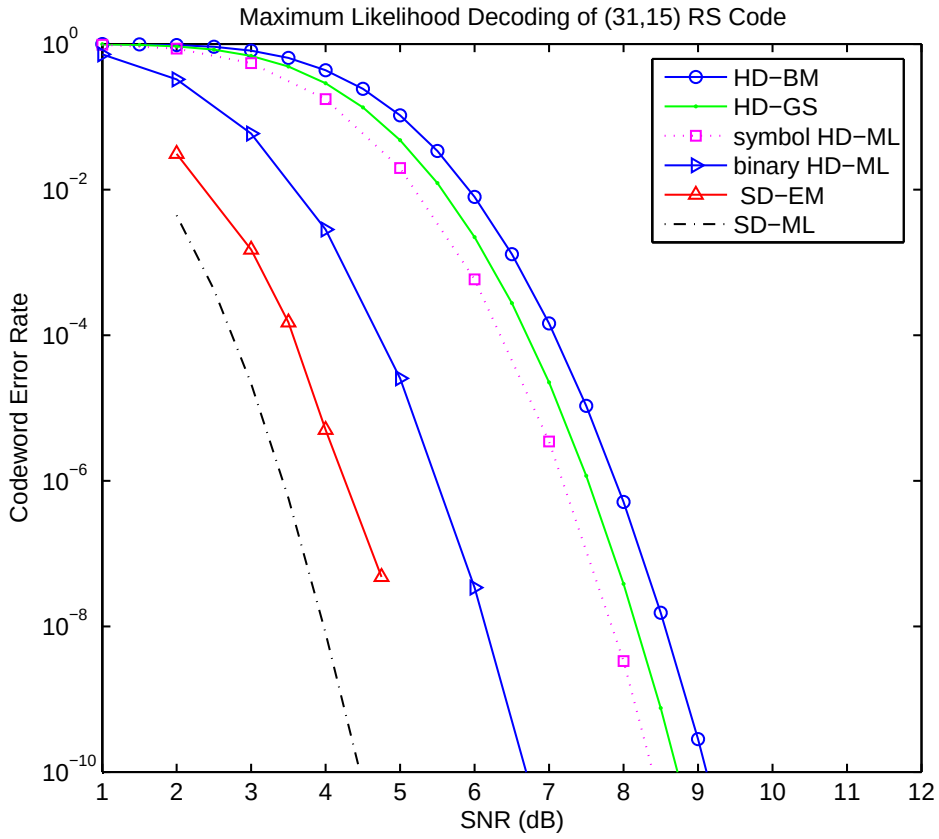

Fig. 10. Performance of the binary image of the $(31, 15)$ RS code over $\mathbb{F}_{32}$ transmitted over AWGN channels. The symbol-level HD-BM and the HD-GS algorithms are compared. Bit-level and symbol-level hard-decision decoders are labeled 'binary HD-ML' and 'symbol HD-ML' respectively. The TSB on the bit-level SD-ML error probability is labeled 'SD-ML' and is compared with the bit-level soft-decision algorithm of [12] labeled 'SD-EM'.

specific combination of binary Hamming weights in the specified partitions. As we saw in the previous section, the binary image is not unique, so we will resort again to an *averaged* binary PWE.

Theorem 19: Let $\mathbb{P}^{\mathcal{T}}(\mathcal{X}_1, \mathcal{X}_2, ..., \mathcal{X}_p)$ *be the partition weight generating function (PWGF) of an* (n, k) *code over* F_{2^m}*, and* $\mathcal{T}_b$ *be the partitioning of the coordinates of* $\mathcal{C}^b$ *induced by* $\mathcal{T}$ *when the symbols in each partition are represented by bits, then the average binary PWGF is*

$$\tilde{\mathbb{P}}^{\mathcal{T}_b}_{\mathcal{C}^b}(\mathcal{Z}_1, \mathcal{Z}_2, ..., \mathcal{Z}_p) = \mathbb{P}^{\mathcal{T}}_{\mathcal{C}}(F(\mathcal{Z}_1), F(\mathcal{Z}_2), ..., F(\mathcal{Z}_p)),$$

where $F(\mathcal{Z}) = \frac{1}{2^m-1}(1 + \mathcal{Z})^m - 1$.

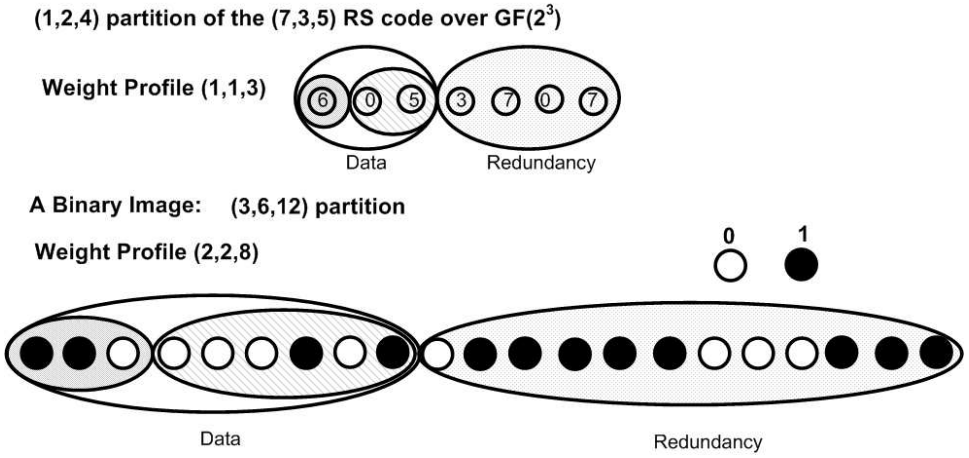

Fig. 11. A codeword in the $(7,3,5)$ RS code over $\mathbb{F}_8$ is shown with a $(1,2,4)$ partition of its coordinates. For a specific binary representation, the binary image is shown with the implied $(3,6,12)$ partition of its coordinates. We emphasize that the weight profile of the binary image is not easily derived from that on the symbol level.

Proof: Assuming a binomial distribution of the bits in a nonzero symbol, the probability that the binary representation of a nonzero symbol has weight i is equal to the coefficient of $\mathcal{Z}^i$ in $\frac{1}{2^m-1}\sum_{i=1}^{m}\binom{m}{i}\mathcal{Z}^i$. If the weight of the jth partition is w_j, then the average binary weight generator function of its binary image is $\left(\frac{1}{2^m-1}\sum_{i=1}^{m}\binom{m}{i}\mathcal{Z}_j^i\right)^{w_j}$ under the assumption that all the non-zero symbols are independent and equally probable. Consider a codeword with a weight profile $(w_1, w_2, ..., w_p)$, then the probability that the weight profile of its binary image is $(b_1, b_2, .., b_p)$ is given by the coefficient of $\mathcal{Z}_1^{b_1}\mathcal{Z}_2^{b_2}...\mathcal{Z}_p^{b_p}$ in $\prod_{j=1}^{p}\left(\frac{1}{2^m-1}\sum_{i=1}^{m}\binom{m}{i}\mathcal{Z}_j^i\right)^{w_j}$. By multiplying with the number of such codewords, $A^{\mathcal{T}}(w_1, w_2, ..., w_p)$, the result follows. ■

For systematic codes, the binary IOWE could be derived from the binary PWE as in (11) (Unless otherwise stated, when speaking of binary weight enumerators of codes over F_{2^m} it is understood that we mean the ensemble average binary weight enumerator.) For example, the coefficient of $\mathcal{X}^w\mathcal{Y}^h$ in $\tilde{\mathbb{P}}(\mathcal{X}\mathcal{Y}, \mathcal{Y}, ..., \mathcal{Y})$ is the number of codewords with input binary weight w in the first partition and a total average binary weight h. In the following corollary, we give a closed form expression for the binary IOWE, $\tilde{O}(w_b, h_b)$.

Corollary 20: Let $O_{\mathcal{C}}(w,h)$ be the input-output weight enumerator of an (n,k,d) code $\mathcal{C}$, defined over $\mathbb{F}_{2^m}$ corresponding to an $(s, n-s)$ partition of its coordinates, then the average

binary IOWE of $\mathcal{C}^b$ is given by

$$\tilde{O}_{\mathcal{C}^b}(w_b, h_b) = \sum_{w=0}^{s}\sum_{h=w}^{n} \frac{O_{\mathcal{C}}(w,h)}{(2^m-1)^h} \left(\sum_{j=0}^{h-w}(-1)^{h-w-j}\binom{h-w}{j}\binom{jm}{h_b-w_b}\right)\left(\sum_{j=0}^{w}(-1)^{w-j}\binom{w}{j}\binom{jm}{w_b}\right)$$

for $h_b \geq d$.

Proof: For the given $(s, n-s)$ partition, the split weight enumerator of $\mathcal{C}$ is $\mathbb{P}_{\mathcal{C}}(\mathcal{X},\mathcal{Y}) = \sum_{w=0}^{s}\sum_{h=w}^{n} O_{\mathcal{C}}(w,h)\mathcal{X}^w\mathcal{Y}^{h-w}$. From the Theorem 19 and (9), $\tilde{O}_{\mathcal{C}^b}(w_b,h_b)$ is the coefficient of $\mathcal{X}^{w_b}\mathcal{Y}^{h_b}$ in

$$\tilde{\mathbb{O}}_{\mathcal{C}^b}(\mathcal{X},\mathcal{Y}) = \frac{1}{(2^m-1)^h}\sum_{w=0}^{s}\sum_{h=w}^{n} O_{\mathcal{C}}(w,h)((1+\mathcal{Y}\mathcal{X})^m-1)^w((1+\mathcal{Y})^m-1)^{h-w}. \tag{65}$$

Since $((1+\mathcal{Y}\mathcal{X})^m-1)^w = \sum_{j=0}^{w}\binom{w}{j}(-1)^{w-j}(\sum_{i=0}^{mj}\binom{mj}{i}\mathcal{X}^i\mathcal{Y}^i)$ and $((1+\mathcal{Y})^m-1)^{h-w} = \sum_{j=0}^{h-w}\binom{h-w}{j}(-1)^{h-w-j}(\sum_{i=0}^{mj}\binom{mj}{i}\mathcal{Y}^i)$, the result follows by substituting in (65). ■

The IOWE of the binary image will be useful in the analysis of the bit error probability of MDS codes when their binary image is transmitted. In Section IV (c.f. Theorem 8), we showed that MDS codes have the multiplicity property. Now, we will show that a binary image of an MDS code with a weight enumerator equal to that of the average binary weight enumerator, if it exists, will also have the multiplicity property.

Theorem 21: Let $\mathcal{C}$ be an (n,k,d) MDS code over $\mathbb{F}_{2^m}$ with the multiplicity property and $\tilde{E}(h_b)$ be the average binary weight enumerator of $\mathcal{C}^b$. If $\tilde{O}(w_b,h_b)$ is the average binary IOWE of $\mathcal{C}^b$, where the partition of the coordinates of $\mathcal{C}^b$ is induced by an $(s,n-s)$ partition of the coordinates of $\mathcal{C}$, then for $h_b \geq d$

$$\frac{\sum_{w_b=1}^{ms} w_b\,\tilde{O}(w_b,h_b)}{m\,s} = \frac{h_b\,\tilde{E}(h_b)}{m\,n}.$$

Proof: We will begin by proving it for the special case of $s=1$. Since $\mathcal{C}$ has property $\mathcal{M}$, then $O(1,h) = \frac{h}{n}E(h)$. It follows from Corollary 20 that

$$\tilde{O}(w_b,h_b) = \binom{m}{w_b}\sum_{h=0}^{n}\frac{h}{n}\frac{E(h)}{(2^m-1)^h}\sum_{j=0}^{h-1}(-1)^{h-1-j}\binom{h-1}{j}\binom{jm}{h_b-w_b}. \tag{66}$$

By changing the order of the summations we have

$$\sum_{w_b=1}^{m} w_b\tilde{O}(w_b,h_b) = \sum_{h=0}^{n}\frac{h}{n}\frac{E(h)}{(2^m-1)^h}\sum_{j=0}^{h-1}(-1)^{h-1-j}\binom{h-1}{j}\sum_{w_b=1}^{m} w_b\binom{m}{w_b}\binom{jm}{h_b-w_b}. \tag{67}$$

By observing that $w_b\binom{m}{w_b} = m\binom{m-1}{w_b-1}$, it follows that the rightmost summation in (68) is equal to $m\sum_{w_b}\binom{m-1}{w_b-1}\binom{mj}{h_b-1-(w_b-1)} = m\binom{m(j+1)-1}{h_b-1}$. By doing a change of variables $\alpha = j+1$ and observing that $\binom{m\alpha-1}{h_b-1} = \frac{h_b}{m\alpha}\binom{m\alpha}{h_b}$ and rearranging it follows that the total weight of m coordinates in the binary image $\mathcal{C}_b$, corresponding to a single coordinate in $\mathcal{C}$, is

$$\begin{aligned}\sum_{w_b=1}^{m} w_b\tilde{O}(w_b,h_b) &= \frac{1}{n}h_b\sum_{h=1}^{n}\frac{E(h)}{(2^m-1)^h}\sum_{\alpha=1}^{h}(-1)^{h-\alpha}\binom{h}{\alpha}\binom{m\alpha}{h_b}\\ &= \frac{h_b}{n}\tilde{E}(h_b). \end{aligned} \tag{68}$$

If the input partition has s coordinates of $\mathcal{C}$, the result follows by summing the weights of the individual coordinates. ■

This means that if the weight of a symbol coordinate is $(h/n)E(h)$ in $\mathcal{C}_h$, then the average weight of its binary image is $(h_b/n)\tilde{E}(h_b)$ in $\mathcal{C}_{h_b}^b$. It will be interesting to determine whether this will still be true for any binary representation. As we will see in the next section, the result of Theorem 21 can simplify the analysis of the bit error probability of MDS codes.

IX. Symbol and Bit Error Probabilities

In section VII, we showed how one can analyze the codeword error probability of various RS code decoders. In this section, we study the symbol and bit error probabilities of systematic MDS codes. In general, systematic coding is preferred over non-systematic coding. It has also been shown that maximum likelihood (ML) decoding of binary linear codes achieves the least bit error probability when the code is systematic [38].

Given a symbol-level decoder (soft-decision or hard-decision decoder), the codeword error error probability (CEP) at a certain signal to noise ratio (SNR) γ will be a function of the SNR γ and the code weight enumerator $E(h)$. In the remaining of this paper, we will denote the CEP at a signal to noise ratio (SNR) γ by $\Phi_c\left(E(h),\gamma\right)$. For linear codes, union upper-bounds on the performance of symbol-based decoders are of the form

$$\Phi_c\left(E(h),\gamma\right) \leq \sum_{h=d}^{n} E(h)\mathcal{U}(\gamma,h) \tag{69}$$

for some function $\mathcal{U}$ of the SNR γ and weight h.

Tighter upper bounds can be of the form

$$\Phi_c\left(E(h),\gamma\right) \leq \min_{\alpha}\left\{\sum_{h=d}^{\alpha} E(h)\mathcal{V}(\gamma,h) \;+\; \mathcal{F}(\gamma,\alpha)\right\} \tag{70}$$

for some functions $\mathcal{V}$ and $\mathcal{F}$ of γ and h. For example, tight upper bounds on the performance of bit-level and symbol-level hard-decision maximum likelihood decoders admit to the above form and are given by [29, Lemma 1] and [33, Theorem 2] respectively. The codeword error probability of the HD Berlekamp-Massey decoder is the probability that the received word lies in the decoding sphere of a codeword other than the transmitted word. It is also determined by the weight enumerator and has the form of the union bound as in (69);

$$\Phi_c\left(E(h),\gamma\right) \leq \sum_{h=d}^{n} E(h) \sum_{t=0}^{\tau} P_t^h(\gamma), \tag{71}$$

where $P_t^h(\gamma)$ is the probability that a received word is exactly Hamming distance t from a codeword of weight h and $\tau = \lfloor (d-1)/2 \rfloor$ is the Hamming decoding radius [28] [32].

Given an upper bound on the CEP of a symbol-based decoder, it is well known that the symbol error probability (SEP) $\Phi_s(\gamma)$ can be derived from the CEP $\Phi_c(\gamma)$ by substituting $E(h)$ with

$$Q(k,h) = \sum_{w=1}^{k} \frac{w}{k}\, O(w,h), \tag{72}$$

(e.g., [32, (10-14)]). From Theorem 8, the common approximation

$$Q(k,h) \approx \frac{h}{n} E(h) \tag{73}$$

is exact for MDS codes and

$$\Phi_s(\gamma) \;\; = \;\; \Phi_c\left(E(h),\gamma\right)\big|_{E(h):=Q(k,h)}\,. \tag{74}$$

In other words, if the CEP is given by (69) or (70), the SEP will be respectively bounded by

$$\Phi_s(\gamma) \;\; \leq \;\; \sum_{h=d}^{n} \frac{h}{n} E(h) \mathcal{U}(\gamma,h), \tag{75}$$

$$\Phi_s(\gamma) \;\; \leq \;\; \min_{\alpha} \left\{ \sum_{h=d}^{\alpha} \frac{h}{n} E(h) \mathcal{V}(\gamma,h) \;+\; \mathcal{F}(\gamma,\alpha) \right\}. \tag{76}$$

In case the binary image of an RS code is transmitted and the decoder is a bit-level decoder, performance analysis of the decoder will utilize the binary weight enumerator of the code. As we discussed in Section VII, the ensemble average binary weight enumerators become handy when analyzing the performance of the binary images of RS codes. As is the case of symbol

based decoders, upper bound on the CEP of bit-level decoders admit the union bound forms

$$\Phi_c\left(\tilde{E}(h),\gamma\right) \leq \sum_{h=d}^{nm}\tilde{E}(h)\Upsilon(\gamma,h) \tag{77}$$

$$\Phi_c\left(\tilde{E}(h),\gamma\right) \leq \min_{\alpha}\left\{\sum_{h=d}^{\alpha}\tilde{E}(h)\mathcal{J}(\gamma,h)\ +\ \mathcal{G}(\gamma,\alpha)\right\} \tag{78}$$

for some functions Υ, $\mathcal{J}$ and $\mathcal{G}$ of the SNR γ and the weight h. For example, the union bounds of SD and HD decoding of (61) and (64) are of the form of (77), whereas the Poltyrev tighter version of these bounds follow the form of (78).

From Theorem 21, we know that for any k (symbol) coordinates of the MDS code

$$\tilde{Q}(mk,h)=\sum_{w=1}^{mk}\frac{w}{mk}\tilde{O}(w,h)=\frac{h}{mn}\tilde{E}(h). \tag{79}$$

It follows that the bit error probability (BEP) can be bounded by (e.g., [21], [39])

$$\Phi_b(\gamma) = \Phi_c\left(\tilde{E}(h),\gamma\right)\Big|_{\tilde{E}(h):=\tilde{Q}(mk,h)} \tag{80}$$

$$\leq \min_{\alpha}\left\{\sum_{h=d}^{\alpha}\frac{h}{mn}\tilde{E}(h)\mathcal{J}(\gamma,h)\ +\ \mathcal{G}(\gamma,\alpha)\right\} \tag{81}$$

$$\leq \sum_{h=d}^{nm}\frac{h}{mn}\tilde{E}(h)\Upsilon(\gamma,h). \tag{82}$$

X. Multiuser Error Probability

We consider the case when a systematic RS code is shared among different users or applications. The systematic symbols are shared among the different users where the coordinates of the code are partitioned according to an $\mathcal{T}:(n_1,n_2,...,n_{p-1},n-k)$ partition. The ith partition of size n_i is assigned to the ith user and the last partition constitutes of the redundancy symbols. Since the considered codes are linear, we assume that the all zero codeword is transmitted. If a codeword of symbol weight h and of weight w_j in the jth partition is erroneously decoded, a fraction $\frac{w_j}{n_j}$ of the jth user's symbols are received in error. It follows that the jth user's symbol error probability could be written as (cf. (87))

$$\Phi_{\mathrm{s}}^{\mathrm{j}}(\gamma)=\Phi_{\mathrm{c}}\left(Q^j(n_j,h),\gamma\right), \tag{83}$$

where

$$Q^j(n_j,h)=\sum_{w=1}^{n_j}\frac{w}{n_j}O^j(w,h) \tag{84}$$

and $O^j(w,h)$ is the jth partition input output weight enumerator derived from the PWE as in (11). The following theorem gives an important result regarding the multiuser error probability of MDS (RS) codes:

Theorem 22: If a systematic linear MDS code is shared among different users, all users have the same unconditional symbol error probability regardless of the sizes of the partitions assigned to them.

Proof: The SEP of a certain user j, whose partition's size is n_j, is given by (83). It is sufficient to show that for two different users i and j with partitions of sizes n_i and n_j respectively, such that $n_i \neq n_j$, $Q^j(n_j,h) = Q^i(n_i,h)$. From Theorem (8), it follows that for an arbitrary partition of size n_j, $Q^j(n_j,h) = \frac{h}{n}E(h)$. Since this result does not depend on the size of the partition nor on the orientation of the coordinates with respect to it, we are done. ■

Now, consider the case when the binary image of an RS code is transmitted and the decoder is a bit-level hard-decision or soft-decision decoder. The systematic coordinates will be partitioned among different users where the partitions on the bit level will follow from the partitions on the symbol level (e.g. Fig. 11). In case of a bit-level decoder, the bit error probability of the jth user can be given by

$$\Phi_b^j(\gamma) = \Phi_c\left(\tilde{Q}^j(mn_j,h),\gamma\right), \tag{85}$$

such that

$$\tilde{Q}^j(mn_j,h) = \sum_{w=1}^{mn_j} \frac{w}{mn_j}\tilde{O}^j(w,h), \tag{86}$$

where $\tilde{O}^j(w,h)$ is the average binary input output weight enumerator of the jth user and $\frac{w}{mn_j}\tilde{O}^j(w,h)$ is the fraction of the jth user's bits received in error when a codeword of total weight h and weight w in the jth partition is erroneously decoded given that the all zero codeword was transmitted.

Theorem 23: For systematic MDS linear codes, the average unconditional bit error probability of all users is the same regardless of the number of symbols in each partition or the orientation of the partition assigned to them.

Proof: Let users i and j be assigned two different partitions of $\mathcal{C}$ with different sizes n_i and n_j. Now consider the binary images of these partitions. Equations (79) and (85) imply that both users have the same average bit error probability. ■

Now that we have shown that the unconditional symbol and bit error probability are the same for all partitions (users) regardless of their size, we can ask questions about the conditional error probability. Using the results in this paper, one could answer interesting questions about the conditional multiuser error probability. Since the code is linear, we will assume that the all-zero codeword is transmitted. For example, the conditional CEP given that for any codeword no more than a fraction p of the jth user's symbols are ever received in error is given by [2]

$$\underline{\Phi_c}(\gamma) \quad = \quad \Phi_c \left(\sum_{w_j=0}^{\lfloor pn_j \rfloor} O^j(w_j, h), \gamma \right) \tag{87}$$

where a hard-decision symbol level decoder with a decoding radius τ was assumed. We only considered error events due to codewords whose weight in the jth partition is not greater than pn_j. Recall that in the unconditional case $\sum_{w_j=0}^{\lfloor pn_j \rfloor} O^j(w_j, h)$ is replaced by $E(h) = \sum_{w_j=0}^{n_j} O^j(w_j, h)$.

Define the following weight enumerator

$$O^{i,j}(w_i, w_j, h) \stackrel{\Delta}{=} |\{\boldsymbol{c} \in \mathcal{C} : (\mathcal{W}(\boldsymbol{c}[N_i]) = w_i) \wedge (\mathcal{W}(\boldsymbol{c}[N_j]) = w_j) \wedge (\mathcal{W}(\boldsymbol{c}) = h)\}|. \tag{88}$$

The conditional CEP given that a codeword error results in all ith user's symbols received correctly while all jth user's symbols received erroneously is given by

$$\underline{\Phi_c}(\gamma) = \Phi_c \left(\sum_{h=d}^{n} O^{i,j}(0, n_j, h), \gamma \right) \tag{89}$$

where assuming that the all-zero codeword is transmitted we only considered codewords with a zero weight in the ith partition and a full weight in the jth partition.

In general, for a p-partition of the coordinates, let $\mathcal{P}$ and $\mathcal{Q}$ be the set of users (partitions) whose symbols are all received correctly and erroneously, respectively, in case of a codeword error. Let $\mathcal{O}$ be the set of users with no condition on their error probability. The conditional error probability is calculated by considering only the codewords which have a full weight for the coordinates in $\mathcal{Q}$ and a zero weight for the coordinates in $\mathcal{P}$. By considering only such combinations in the sum of (5), the conditional PWGF is derived as

$$\underline{\mathbb{P}}(\mathcal{X}_1, \mathcal{X}_2, ..., \mathcal{X}_p) = \sum_{i \in \Delta} \sum_{w_i=0}^{n_i} A(w_1, w_2, ..., w_p) \mathcal{X}_1^{w_1} \mathcal{X}_2^{w_2} ... \mathcal{X}_p^{w_p} \left| \begin{array}{ll} w_i = 0, & \text{if } i \in \mathcal{P}; \\ w_i = n_i, & \text{if } i \in \mathcal{Q}. \end{array} \right. . \tag{90}$$

[2]Conditional functions will have have the same notation as the unconditional ones except for an underbar.

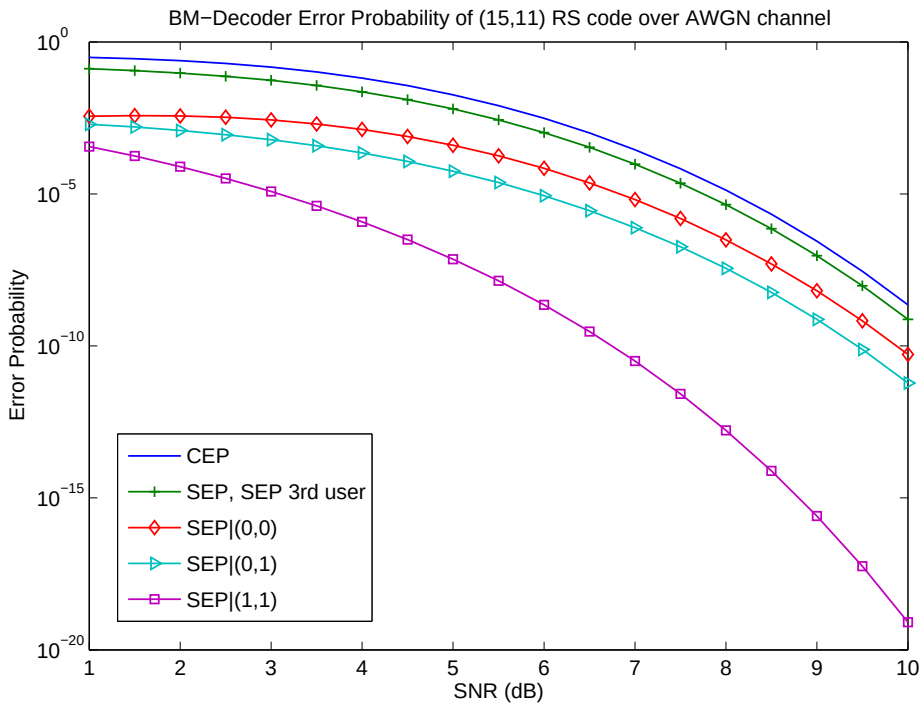

Fig. 12. Conditional multiuser decoder error probability of the Berlekamp-Massey decoder of Example 4. The unconditional CEP and SEP are labeled 'CEP' and 'SEP' respectively. The conditional SEP of the third user for cases 1, 2 and 3 are labeled 'SEP$|(0,0)$', 'SEP$|(0,1)$' and 'SEP$|(1,1,)$' respectively.

The conditional symbol error probability of the jth user is

$$\underline{\Phi_s^j}(\gamma) = \Phi_c\left(\underline{Q^j(k,h)},\gamma\right), \tag{91}$$

where $\underline{Q^j(k,h)} = \sum_{w=1}^{n_j} \frac{w}{n_j}\underline{O^j}(w,h)$ and $\underline{O^j}(w,h)$ is the conditional IOWE of the jth partition and is derived from $\underline{\mathbb{P}}(\mathcal{X}_1,\mathcal{X}_2,...,\mathcal{X}_p)$ (see (10)). For example, if the first partition contains header information, then the conditional symbol error probability of the ith user given that the header is received correctly can be calculated by

$$\underline{\Phi_s^j}(\gamma) = \Phi_c\left(\sum_{w=1}^{n_j}\frac{w}{n_j}O^{1,j}(0,w,h),\gamma\right). \tag{92}$$

Similarly, for bit-level decoding of the code's binary image, $\underline{\tilde{Q}^j(mk,h)}$ will be derived from $\underline{\tilde{\mathbb{P}}}(\mathcal{X}_1,\mathcal{X}_2,...,\mathcal{X}_p)$. If the users in $\mathcal{P}$ and $\mathcal{Q}$ have zero and one bit error probability respectively, the conditional binary PWGF only takes into account such codewords that have a zero binary

weight for the partitions in $\mathcal{P}$ and a full binary Hamming weight for the partitions in $\mathcal{Q}$. The conditional BEP of the jth user follows by the substitution $\tilde{E}(h) := \underline{\tilde{Q}^j(mk,h)}$ in (80).

Example 4: Consider an systematic $(15,11,5)$ RS code and a partition $\mathcal{T} = (3,3,5,4)$ of its coordinates where the last partition has the redundancy symbols and each of the first three partitions is assigned to a different user. The first partition may be assigned to be the header. Let the RS code be transmitted over an AWGN channel and decoded by a hard-decision bounded minimum distance (Berlekamp-Massey) decoder. From (71), (71), (87) and Theorem 22 it follows that the CEP and SEP of any user is equal to the overall SEP and can be expressed as, respectively,

$$\begin{aligned}\Phi_c(\gamma) &= \sum_{h=5}^{15} E(h) \sum_{t=0}^{\tau} P_t^h(\gamma),\\ \Phi_s(\gamma) &= \sum_{h=5}^{15} \frac{h}{15} E(h) \sum_{t=0}^{\tau} P_t^h,\end{aligned}$$

such that $E(h)$ is the weight enumerator as given by (13). The partition weight generating function is given by

$$\mathbb{P}(\mathcal{W},\mathcal{X},\mathcal{Y},\mathcal{Z}) = \sum_{w_1=0}^{3}\sum_{w_2=0}^{3}\sum_{w_3=0}^{5}\sum_{w_4=0}^{4} A^{\mathcal{T}}(w_1,w_2,w_3,w_4)\mathcal{W}^{w_1}\mathcal{X}^{w_2}\mathcal{Y}^{w_3}\mathcal{Z}^{w_4},$$

and the IOWGF of the third user is $\mathbb{O}^3(\mathcal{X},\mathcal{Y}) = \mathbb{P}(\mathcal{X},\mathcal{X},\mathcal{X}\mathcal{Y},\mathcal{X})$. We will now calculate the conditional symbol error probability of the third user under different scenarios.

Case 1: The first two users have a zero error probability. Thus the PWGF conditioned on that the first two partitions have zero weight is

$$\underline{\mathbb{P}}_{(0,0)}(\mathcal{Y},\mathcal{Z}) = \sum_{w_3=0}^{5}\sum_{w_4=0}^{4} A^{\mathcal{T}}(0,0,w_3,w_4)\mathcal{Y}^{w_3}\mathcal{Z}^{w_4}.$$

The conditional IOWGF of the 3rd user is

$$\underline{\mathbb{O}}^3_{(0,0)}(\mathcal{X},\mathcal{Y}) = \underline{\mathbb{P}}_{(0,0)}(\mathcal{X}\mathcal{Y},\mathcal{Y}) = \sum_{w}\sum_{h} \underline{O}^{1,2,3}(0,0,w,h)\mathcal{X}^w\mathcal{Y}^h,$$

It follows that the SEP of the 3rd user conditioned on that the first two users have a zero error probability is

$$\underline{\Phi_s^3}(\gamma) = \sum_{h=d}^{n}\sum_{w=1}^{5} \frac{w}{5}\underline{O}^{1,2,3}(0,0,w,j)\sum_{t=0}^{\tau} P_t^h.$$

Case 2: The first and second users have an SEP of zero and one respectively. The corresponding conditional PWGF is

$$\underline{\mathbb{P}}_{(0,1)}(\mathcal{X},\mathcal{Y},\mathcal{Z}) = \sum_{w_3=0}^{5}\sum_{w_4=0}^{4} A^{\mathcal{T}}(0,3,w_3,w_4)\mathcal{X}^3\mathcal{Y}^{w_3}\mathcal{Z}^{w_4}.$$

The corresponding IOWGF of the 3rd user is

$$\underline{\mathbb{O}}^3_{(0,1)}(\mathcal{X},\mathcal{Y}) = \underline{\mathbb{P}}_{(0,1)}(\mathcal{Y},\mathcal{X}\mathcal{Y},\mathcal{Y}) = \sum_{w}\sum_{h}\underline{O}^{1,2,3}(0,3,w,h)\mathcal{X}^w\mathcal{Y}^h.$$

To calculate the conditional SEP, we proceed as in the previous case.

Case 3: Both the first and second users have an SEP of one. The conditional SEP of the third user is

$$\underline{\Phi}^3_s(\gamma) = \sum_{h=d}^{n}\sum_{w=1}^{5}\frac{w}{5}\underline{O}^{1,2,3}(3,3,w,j)\sum_{t=0}^{\tau}P_t^h.$$

where $\underline{O}^{1,2,3}(3,3,w,h)$ is the coefficient of $\mathcal{X}^w\mathcal{Y}^h$ in $\underline{\mathbb{O}}^3_{(1,1)}(\mathcal{X},\mathcal{Y}) = \underline{\mathbb{P}}_{(1,1)}(\mathcal{Y},\mathcal{Y},\mathcal{X}\mathcal{Y},\mathcal{Y})$ and

$$\underline{\mathbb{P}}_{(1,1)}(\mathcal{W},\mathcal{X},\mathcal{Y},\mathcal{Z}) = \sum_{w_3=0}^{5}\sum_{w_4=0}^{4} A^{\mathcal{T}}(3,3,w_3,w_4)\mathcal{W}^3\mathcal{X}^3\mathcal{Y}^{w_3}\mathcal{Z}^{w_4}.$$

For an AWGN channel and a Berlekamp-Massey decoder, the codeword error probability, symbol error probability and the conditional symbol error probabilities for the third user for the three cases are plotted in Fig. 12. It is observed that the conditional error probability of the third user given that other users have an error probability of one (Case 3) is the lowest compared to the other two cases. The reason is that in Case 3, one only considers errors due to the received word falling closer to codewords at a much larger Hamming distance from the transmitted one, and such an event happens with relatively lower probability. □

The same technique can be used to bound the performance of other symbol based decoders, such as the hard-decision maximum likelihood decoder, under various scenarios. Next we consider analyzing the multiuser error probability when the decoder is a bit level decoder.

Example 5: Consider the $(15,11,5)$ code over $\mathbb{F}_{16}$ partitioned as in Example 4 and an SD bit-level ML decoder is employed at the output of an AWGN channel. The unconditional CEP and BEP are given by, respectively,

$$\begin{aligned}\Phi_c\left(\tilde{E}(h),\gamma\right) &\leq \min_{\alpha}\left\{\sum_{h=5}^{\alpha}\tilde{E}(h)\mathcal{J}(\gamma,h) + \mathcal{G}(\gamma,\alpha)\right\}\\ \Phi_b(\gamma) &= \min_{\alpha}\left\{\sum_{h=5}^{\alpha}\frac{h}{60}\tilde{E}(h)\mathcal{J}(\gamma,h) + \mathcal{G}(\gamma,\alpha)\right\},\end{aligned}$$

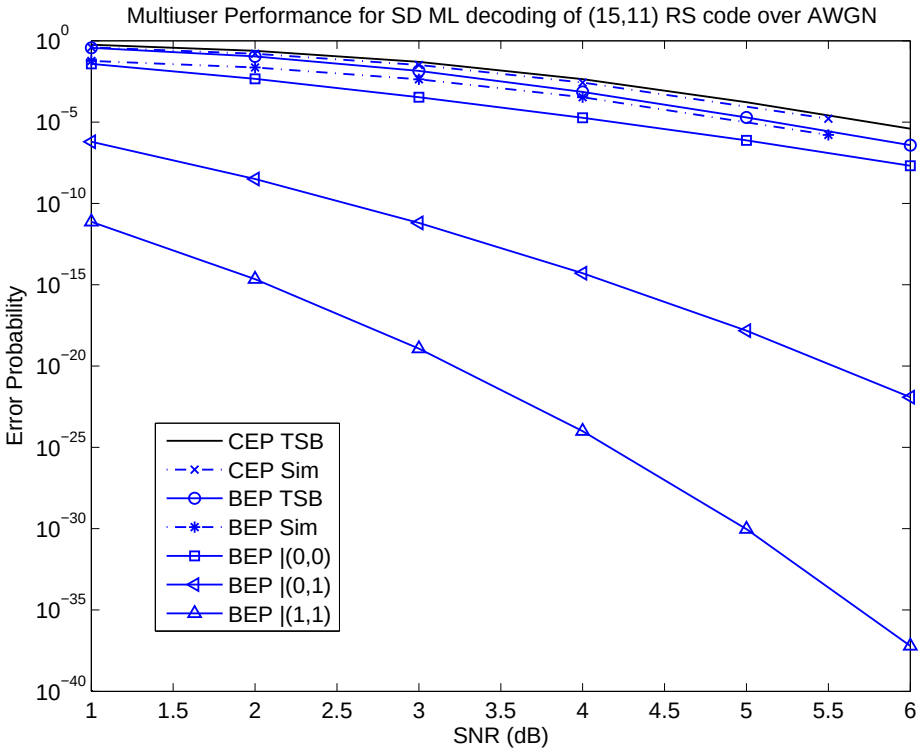

Fig. 13. Conditional multiuser error probability of the bit-level soft-decision maximum-likelihood decoder of Example 5. The conditional bit error probability of cases 1, 2 and 3 are labeled 'BEP$|(0,0)$, 'BEP$|(0,1)$' and 'BEP$|(1,1)$'. The bounds on the unconditional CEP and BEP labeled 'CEP TSB' and 'BEP TSB' are compared with the corresponding simulations labeled 'CEP Sim' and 'BEP Sim'respectively.

where $\mathcal{J}(\gamma, h)$ and $\mathcal{G}(\gamma, \alpha)$ will be determined by the Poltyrev tangential sphere bound [29]. We will now discuss the conditional bit error probability for different cases (as in Example 4):

Case 1: The first two users have a zero error probability. The average binary IOWE of the third user given the first two partitions have a zero weight is

$$\underline{\tilde{\mathbb{O}}}^{3}_{(0,0)}(\mathcal{X}, \mathcal{Y}) = \underline{\tilde{\mathbb{P}}}_{(0,0)}(\mathcal{X}\mathcal{Y}, \mathcal{Y}) = \sum_{h=0}^{60} \sum_{w=0}^{20} \underline{\tilde{O}}^{1,2,3}(0, 0, w, h) \mathcal{X}^w \mathcal{Y}^h,$$

such that $\underline{\tilde{\mathbb{P}}}_{(0,0)}(\mathcal{X}, \mathcal{Y}) = \underline{\mathbb{P}}_{(0,0)}(F(\mathcal{X}), F(\mathcal{Y}))$, and $F(\mathcal{X})$ is as defined in Theorem 19. The conditional BEP of the third user is given by

$$\underline{\Phi}^{3}_{b}(\gamma) = \min_{\alpha} \left\{ \sum_{h=5}^{\alpha} \sum_{w=1}^{20} \frac{w}{20} \underline{\tilde{O}}^{1,2,3}(0, 0, w, h) \mathcal{J}(\gamma, h) + \mathcal{G}(\gamma, \alpha) \right\}.$$

Case 2: The first and second users have a zero and one bit error probability respectively. Let $\tilde{\mathbb{P}}(\mathcal{W},\mathcal{X},\mathcal{Y},\mathcal{Z}) = \mathbb{P}(F(\mathcal{W}),F(\mathcal{X}),F(\mathcal{Y}),F(\mathcal{Z}))$ be the average binary PWGF then

$$\underline{\tilde{\mathbb{P}}}_{(0,1)}(\mathcal{X},\mathcal{Y},\mathcal{Z}) = \text{Coeff}\left(\tilde{\mathbb{P}}(\mathcal{W},\mathcal{X},\mathcal{Y},\mathcal{Z}),\mathcal{W}^0\mathcal{X}^{12}\right)\mathcal{X}^{12}$$

and the conditional IOWE of the third user is

$$\underline{\tilde{O}}^{1,2,3}(0,12,w,h) = \text{Coeff}\left(\tilde{\mathbb{P}}_{(0,1)}(\mathcal{Y},\mathcal{X}\mathcal{Y},\mathcal{Y}),\mathcal{X}^w\mathcal{Y}^h\right).$$

The conditional BEP is then given by

$$\underline{\Phi_b^3}(\gamma) = \min_{\alpha}\left\{\sum_{h=5}^{\alpha}\sum_{w=1}^{20}\frac{w}{20}\underline{\tilde{O}}^{1,2,3}(0,12,w,h)\mathcal{J}(\gamma,h) + \mathcal{G}(\gamma,\alpha)\right\}.$$

Case 3: The average BEP of the first two users is one. In this case, the conditional PWGF can be calculated by

$$\underline{\tilde{\mathbb{P}}}_{(1,1)}(\mathcal{W},\mathcal{X},\mathcal{Y},\mathcal{Z}) = \text{Coeff}\left(\tilde{\mathbb{P}}(\mathcal{W},\mathcal{X},\mathcal{Y},\mathcal{Z}),\mathcal{W}^{12}\mathcal{X}^{12}\right)\mathcal{W}^{12}\mathcal{X}^{12}.$$

One can then proceed to calculate the conditional IOWE and BPE of the third user by

$$\underline{\tilde{O}}^{1,2,3}(12,12,w,h) = \text{Coeff}\left(\tilde{\mathbb{P}}_{(1,1)}(\mathcal{Y},\mathcal{Y},\mathcal{X}\mathcal{Y},\mathcal{Y}),\mathcal{X}^w\mathcal{Y}^h\right)$$

$$\underline{\Phi_b^3}(\gamma) = \min_{\alpha}\left\{\sum_{h=5}^{\alpha}\sum_{w=1}^{20}\frac{w}{20}\underline{\tilde{O}}^{1,2,3}(12,12,w,h)\mathcal{J}(\gamma,h) + \mathcal{G}(\gamma,\alpha)\right\}.$$

In Fig. 13, the TSB on the codeword and bit error probability are plotted and compared to simulations of the ML decoder for a specific basis representation of the RS code. The conditional BEP of the third user is plotted for cases $1,2$ and 3 . As in the previous example, it is observed that the conditional error probability of specific users given that some users have a high error probability decreases with the number of such users. □

Example 6: Consider an systematic $(31,15,17)$ RS code over $\mathbb{F}_{32}$ and a partition $\mathcal{T} = (3,6,6,16)$ of its coordinates where the last partition has the redundancy symbols and each of the first three partitions is assigned to a different user. The first partition may be assigned to be the header. Let the binary image of a RS code be transmitted over an AWGN channel and decoded by a hard-decision symbol-based maximum likelihood decoder decoder. We used the upper bound by El-Khamy *et. al* to bound the performance of the HD-ML decoder over $\mathbb{F}_{32}$ [33]. The CEP, SEP and conditional SEP are of the form of (70), (78) and (91). We consider three cases:

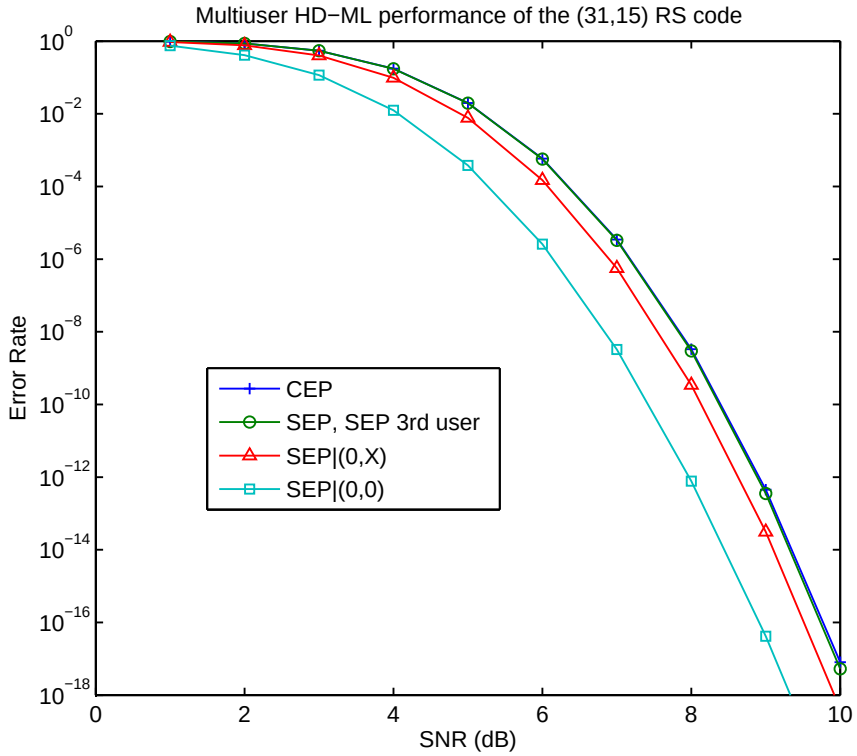

Fig. 14. Conditional multiuser error probability of the symbol-level hard-decision maximum-likelihood decoder of the $(31, 15)$ RS over $\mathbb{F}_{32}$ of Example 6. The unconditional CEP and SEP are plotted (Case 1). The conditional SEP of Cases 2 and 3 are labeled SEP$|(0, X)$ and SEP$|(0, 0)$ respectively.

Case 1: The unconditional error probability of the third user.

Case 2: The symbol error probability of the third user given that the first user (header) is received correctly.

Case 3: The symbol error probability of the third user given that the first two users have their symbols received correctly.

The numerical results are shown in Fig. 14. We observe that the unconditional CEP and SEP are very close. As more and more conditions are imposed, the conditional error probability of the third user decreases. *Case 2*, is of special interest, since in some cases the header will contain the routing information and it will be essential to estimate the error probability in case the information is routed correctly. □

XI. Conclusion

An averaged binary weight enumerator for RS codes is derived and shown to closely estimate an exact one for a specific basis representation. Moreover, it has been shown that as the code length and the field size tend to infinity, the weight enumerator of the ensemble of binary images of Reed-Solomon codes approach that of a random code with the same dimensions. Bounds on the average binary minimum distance were derived. It was thus shown that on average, the ensemble of binary images of RS codes asymptotically satisfy the GV bound. The question remains open, if there exists a specific code in the ensemble that asymptotically satisfies the GV bound. Aided with the ensemble weight enumerator, one can derive tight bounds on the performance of bit-level maximum likelihood decoders. By comparing with simulations, it has been shown, that at least for the $(15, 11)$ RS code, the tangential sphere bound when combined with the ensemble weight enumerator is tight. When proposing new algorithms for decoding RS codes, it is not only important to compare its performance with other algorithms in the literature, but it is also more important to compare its performance with that of other maximum likelihood decoders using the results in this paper. A closed form formula for the partition weight enumerator of maximum distance separable (MDS) codes is derived. The average PWE is derived for the binary image of MDS codes defined over a field of characteristic two. We show that for MDS codes, all the coordinates have the same weight in the subcode composed of codewords with equal weight. We prove that a code has this property iff its dual code has this property. Consequently, it is shown that the first order Reed Muller codes and the extended Hamming codes have this property. A common approximation used to evaluate the symbol and bit error probabilities is thus shown to be exact for MDS codes. These results are employed to study the error probability when a Reed-Solomon code is used in a network scenario and is shared among different users. We show that MDS (e.g. RS) codes have many attractive features which makes their use in networks attractive. It is proved that the unconditional error probability of all the users will be the same regardless of the size of their partitions. As for the conditional error probabilities, they can be a useful measure in determining the performance of a user, if its performance depends on the correct transmission of a certain packet or header.

REFERENCES

[1] M. El-Khamy and R. J. McEliece, "Bounds on the average binary minimum distance and the maximum likelihood performance of Reed Solomon codes," in *42nd Allerton Conf. on Communication, Control and Computing*, 2004.

[2] ——, "The partition weight enumerator of MDS codes and its applications." in *2005 IEEE International Symposium on Information Theory, Adelaide, Australia*, Sept 2005, pp. 926–930.

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

[4] S. B. Wicker and M. J. Bartz, "Type-II hybrid- ARQ protocols using punctured MDS codes," *IEEE Trans. Commun.*, vol. 42, pp. 1431–1440, FEB/MAR/APR 1994.

[5] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, "Network coding theory: Single sources," *Foundations and Trends in Communications and Information Theory*, vol. 2, June 2005.

[6] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *J. Soc. Industrial Appl. Math*, vol. 8, pp. 300–304, 1960.

[7] Berlekamp, R. McEliece, and H. van Tilborg, "On the inherent intractability of certain coding problems," *IEEE Trans. Inform. Theory*, vol. 24, pp. 384–386, May 1978.

[8] V. Guruswami and M. Sudan, "Improved decoding of Reed-Solomon codes and algebraic geometry codes," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 1757–1767, Sept. 1999.

[9] R. J. McEliece, "The Guruswami-Sudan decoding algorithm for Reed-Solomon codes," IPN Progress Report, Tech. Rep. 42–153, May 15 2003.

[10] R. Kötter and A. Vardy, "Algebraic soft-decision decoding of Reed-Solomon codes," *IEEE Trans. Inform. Theory*, vol. 49, no. 11, pp. 2809–2825, Nov. 2003.

[11] M. El-Khamy and R. J. McEliece, "Interpolation multiplicity assignment algorithms for algebraic soft-decision decoding of Reed-Solomon codes," *AMS-DIMACS volume on Algebraic Coding Theory and Information Theory*, vol. 68, 2005.

[12] ——, "Iterative algebraic soft-decision list decoding of Reed-Solomon codes," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 481–490, March 2006.

[13] J. Jiang and K. R. Narayanan, "Iterative soft-input soft-output decoding of Reed-Solomon codes by adapting the parity-check matrix," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3746–3756, Aug. 2006.

[14] I. Sason and S. Shamai, "Performance analysis of linear codes under maximum-likelihood decoding: a tutorial," *Foundations and Trends in Communications and Information Theory*, vol. 3, July 2006.

[15] T. Kasami and S. Lin, "The binary weight distribution of the extended $(2^m, 2^m - 4)$ code of the Reed Solomon code over GF(2^m) with generator polynomial $(x - \alpha)(x - \alpha^2)(x - \alpha^3)$." *Linear Algebra Appl.*, pp. 291–307, 1988.

[16] I. Blake and K. Kith, "On the complete weight enumerator of Reed-Solomon codes." *SIAM J. Disc. Math.*, vol. 4, no. 2, pp. 164–171, May 1991.

[17] C. Retter, "The average binary weight enumerator for a class of generalized Reed-Solomon codes," *IEEE Trans. Inform. Theory*, vol. 37, no. 2, pp. 346–349, March 1991.

[18] S. Dolinar, D. Divsalar, and F. Pollara, "Code performance as a function of block size," TMO Progress Report, Tech. Rep. 42-133, 1998.

[19] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[20] T. Kasami, T. Takata, K. Yamachita, T. Fujiwara, and S. Lin, "On bit error probability of a concatenated coding scheme." *IEEE Trans. Commun.*, vol. 45, no. 5, pp. 536–543, May 1997.

[21] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, “Serial concatenation of interleaved codes: Performance analysis, design and iterative decoding.” *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 909–926, May 1998.

[22] R. J. McEliece, *The Theory of Information and Coding*, 2nd ed. Cambridge: Cambridge U. Press, 2002.

[23] J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, 2nd ed. Cambridge: Cambridge U. Press, 2001.

[24] D. Torrieri, “Information-bit, information-symbol, and decoded-symbol error rates for linear block codes,” *IEEE Trans. Commun.*, pp. 613–617, May 1988.

[25] M. El-Khamy, “The average weight enumerator and the maximum likelihood performance of product codes,” in *International Conference on Wireless Networks, Communications and Mobile Computing, WirelessCom Information Theory Symposium, Hawaii*, vol. 2, June 2005, pp. 1587–1592.

[26] F. Chiaraluce and R. Garello, “Extended Hamming product codes analytical performance evaluation for low error rate applications,” *IEEE Transactions on Wireless Communications*, vol. 3, pp. 2353–2361, Nov. 2004.

[27] M. El-Khamy and R. Garello, “On the weight enumerator and the maximum likelihood performance of linear product codes,” submitted to IEEE Trans. on Information Theory, Dec. 2005.

[28] R. J. McEliece and L. Swanson, “On the decoder error probability of Reed-Solomon codes,” *IEEE Trans. Inform. Theory*, vol. 32, no. 5, pp. 701–703, Sep. 1986.

[29] G. Poltyrev, “Bounds on the decoding error probability of binary linear codes via their spectra,” *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1284–1292, July 1994.

[30] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2001.

[31] D. Divsalar, “A simple tight bound on error probability of block codes with application to turbo codes,” TMO Progress Report, NASA, JPL, Tech. Rep. 42–139, 1999.

[32] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice Hall, 1995.

[33] M. El-Khamy, H. Vikalo, B. Hassibi, and R. J. McEliece, “On the performance of sphere decoding of block codes,” in *2006 IEEE International Symposium on Information Theory, Seattle, Washington*, June 2006.

[34] ——, “Performance of sphere decoding of block codes,” submitted to IEEE Transactions on Communications, Feb. 2006.

[35] L. Bahl, J. Cocke, F. Jeinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate.” *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–7, Mar 1974.

[36] M. Kan, private Communication.

[37] M. Fossorier, “Critical point for maximum likelihood decoding of linear block codes,” *IEEE Commun. Lett.*, vol. 9, no. 9, 2005.

[38] M. Fossorier, S. Lin, and D. Rhee, “Bit-error probability for maximum-likelihood decoding of linear block codes and related soft-decision decoding methods,” *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 3083–3090.

[39] I. Sason and S. Shamai, “Improved upper bounds on the ML decoding error probability of parallel and serial concatenated turbo codes via their ensemble distance spectrum.” *IEEE Trans. Inform. Theory*, vol. 46, no. 1, pp. 24–47, Jan 2000.