On ML Redundancy of Codes

Junsheng Han[1]
Qualcomm Incorporated
San Diego, CA 92121, USA
Email: jhan@qualcomm.com

Paul H. Siegel
Center for Magnetic Recording Research, UCSD
La Jolla, CA 92093-0401, USA
Email: psiegel@ucsd.edu

***Abstract*— The *ML redundancy* of a code is defined as the smallest number of rows in its parity-check matrix such that a message-passing decoder working in the corresponding Tanner graph achieves maximum-likelihood (ML) performance on an erasure channel. General upper bounds on ML redundancy are obtained. In particular, it is shown that the ML redundancy of a q-ary code is at most the number of minimal codewords in its dual code, divided by $q-1$. Special upper bounds are derived for codes whose dual code contains a covering design. For example, the ML redundancy of a Simplex code of length n is shown to be no greater than $(n^2-4n+9)/6$.**

I. Introduction

In this paper, a *code* is a linear block code, which is a k-dimensional linear subspace of $\mathbb{F}_q^n$, where $\mathbb{F}_q$ is a Galois field with q elements. Such a code is usually conveniently referred to as a $[n,k]_q$ code. At times, we will also write $[n,k,d]_q$ to signify that the code has minimum Hamming distance d. A code is usually defined as the row space of a *generator matrix*, G, or, equivalently, the kernel of a *parity-check matrix*[2], H. The dimension of the row space of H is the *redundancy* of the code, denoted by $r = n - k$.

Let H be a parity-check matrix for code $\mathcal{C}$. The *iterative H-decoder* is the message-passing erasure decoder corresponding to the Tanner graph defined by H. (We shall assume the channel is an erasure channel throughout the paper.) An iterative H-decoder works effectively by finding, at each iteration, a parity-check equation that involves a single erasure, solving for the erasure, and repeating the process until either all erasures have been decoded, or no such equations can be found (causing decoding failure). We are interested in the case where H is chosen such that the iterative H-decoder fails only if the same set of erasures cannot be uniquely decoded by even the word maximum-likelihood (ML) decoder, and further in the minimum number of rows in H for this to be possible.

Let $[n] \stackrel{\text{def}}{=} \{1,2,\ldots,n\}$. The *support* of an n-dimensional vector $\boldsymbol{x}$, denoted by $\mathrm{supp}(\boldsymbol{x})$, is the set of coordinates where $\boldsymbol{x}$ is not zero, i.e. $\mathrm{supp}(\boldsymbol{x}) \stackrel{\text{def}}{=} \{i \in [n] : x_i \neq 0\}$. For an n-column parity-check matrix H and $S \subseteq [n]$, let $H(S)$ denote the submatrix of H consisting of the columns indexed by S. For a given H, a *stopping set* is any set $S \subseteq [n]$, such that $H(S)$ does not contain a row of weight one. An *erasure pattern*, a set of code coordinates that are erased, can be corrected by the ML decoder if and only if it does not contain the support of a codeword, in which case we say that the erasure pattern is *correctable*. In comparison, for the iterative H-decoder, an erasure pattern can be corrected if and only if it does not contain a stopping set. Hence, the iterative H-decoder is ML if and only if none of the correctable erasure patterns contains a stopping set. Since any subset of a correctable erasure pattern is still correctable, this is equivalent to the condition that the set of correctable erasure patterns and the set of stopping sets be disjoint. For the sake of discussion, we say that a matrix M *covers* S if $M(S)$ contains a row of weight one. This definition applies to the special case where M is a vector. Clearly, the iterative H-decoder achieves ML performance if and only if all correctable erasure patterns are covered by H.

We define the *ML redundancy* of code $\mathcal{C}$, denoted by $\gamma(\mathcal{C})$, as the smallest number of rows in a parity-check matrix H such that an iterative H-decoder for $\mathcal{C}$ exists that achieves ML performance. In this paper, we will develop bounds on $\gamma(\mathcal{C})$ both in general and for certain classes of codes.

Let H^* denote a matrix consisting of all codewords of $C^\perp$ as rows. The following lemma is a known result [1], which shows that $\gamma(\mathcal{C})$ is well-defined, i.e. it is always possible to achieve ML performance with an iterative decoder on the erasure channel.

Lemma 1 *The iterative H^*-decoder is ML.*

Proof: If $S \subseteq [n]$ is correctable, then for any parity-check matrix H, the column vectors of $H(S)$ are linearly independent. Hence, the rows of $H(S)$ contain a basis of $\mathbb{F}_q^{|S|}$. Therefore, all possible q-ary $|S|$-tuples appear as rows in $H^*(S)$ the same number of times. In particular, $H^*(S)$ must contain a row of weight one. ■

The next lemma shows that if a matrix H has all its rows taken from $\mathcal{C}^\perp$ and covers all correctable erasure patterns for $\mathcal{C}$, then it must indeed be a parity-check matrix for $\mathcal{C}$. Hence, to verify that an iterative H-decoder achieves ML performance, it suffices to verify that H covers all correctable erasure patterns.

Lemma 2 *If $\mathcal{C} \subseteq \mathrm{Null}(H)$ and H covers all correctable erasure patterns for $\mathcal{C}$, then $\mathcal{C} = \mathrm{Null}(H)$, where $\mathrm{Null}(H)$ is the null space of H.*

Proof: The redundancy of $\mathcal{C}$ being r implies that correctable erasure patterns of size r exist. Let S be such an erasure pattern. Since S is covered by H, it is covered by some

[1]This work was done while the first author was at the Center for Magnetic Recording Research at the University of California, San Diego.

[2]For us, a parity-check matrix may consist of rows that are linearly dependent, as long as its kernel defines the code.

978-1-4244-2571-6/08/$25.00 ©2008 IEEE

row of H, say $\boldsymbol{h}_1$. Let s_1 denote the coordinate that $\mathrm{supp}(\boldsymbol{h}_1)$ and S have in common. Then $S \setminus \{s_1\}$ is a correctable erasure pattern of size $r-1$. And so it must be covered by some other row of H, say $\boldsymbol{h}_2$. Repeating the above argument shows that H contains an $r \times r$ lower triangular matrix, up to row and column permutations. Hence, $\mathrm{rank}(H) = r$, and since $\mathcal{C} \subseteq \mathrm{Null}(H)$, it follows that $\mathcal{C} = \mathrm{Null}(H)$. ∎

The following result can be obtained as an easy extension of [2, Theorem 2].

Theorem 3 *Let $\mathcal{C}$ be a binary code with redundancy r. Then*

$$\gamma(\mathcal{C}) \leq 2^{r-1}. \tag{1}$$

Proof: Omitted. ∎

The same result may be inferred from the work by Hollmann and Tolhuizen [3], who considered general constructions of parity-check collections for binary codes that correct all correctable erasure patterns up to a given size. The results in [3] also showed that equality is achieved in (1) for binary Hamming codes.

Theorem 4 ([3]) *Let $\mathcal{H}_m$ denote the $[2^m-1, 2^m-m-1, 3]_2$ Hamming code. Then*

$$\gamma(\mathcal{H}_m) = 2^{m-1}.$$

Our work was motivated by the recent introduction [4] of stopping redundancy by Schwartz and Vardy. The *stopping redundancy* of a code $\mathcal{C}$, denoted by $\rho(\mathcal{C})$, is the smallest number of rows in a parity-check matrix H, such that the size of the smallest non-empty stopping set(s) (the *stopping distance*) is equal to the minimum distance of $\mathcal{C}$. Stopping redundancy was further studied in [2],[5],[6],[7], and has been extended in a number of interesting ways. In [7], a similar definition was applied to *trapping sets*. In [8], the *stopping redundancy hierarchy* was defined and studied. Recently, the concept was extended to the binary symmetric channel (BSC), where the *pseudoweight redundancy* was introduced [9].

The fundamental observation that underlies all aforementioned work is that the *representation* of a code has an impact on its performance under iterative decoding. In particular, by allowing a more complex representation, the performance of iterative decoding is potentially improved. The definition of stopping redundancy highlights a particular point in the complexity-performance tradeoff. Stopping redundancy hierarchy, on the other hand, aims at a possibly more modest performance target. In this sense, ML redundancy can be viewed as a variation of the concept of stopping redundancy with a more ambitious performance goal. Unlike stopping redundancy, whose definition is appropriate when the channel erasure probability is small (approaches zero), ML redundancy is relevant in general. ML redundancy is also important because it provides a perspective for understanding the relation/difference between ML and iterative decoding.

Work related to our study is also found in [10].

The rest of the paper is arranged as follows. Section II concerns general upper bounds on ML redundancy. In Section III, more specific results are given for codes whose dual code (or rather, the supports of the codewords therein) contains combinatorial designs. Section IV concludes the paper.

II. Upper Bounds

First, note that Theorem 3 can be extended to q-ary codes using ideas similar to those in [3].

Theorem 5 *Let $\mathcal{C}$ be a q-ary code with redundancy $r > 0$. Then*

$$\gamma(\mathcal{C}) \leq q^{r-1}.$$

Proof: Let H be an $r \times n$ parity-check matrix for $\mathcal{C}$, where n is the length of $\mathcal{C}$. Let the row vectors of H be denoted by $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_r$. Define $V \stackrel{\mathrm{def}}{=} \big\{\boldsymbol{h}_1 + \boldsymbol{v} : \boldsymbol{v} \in \mathrm{span}(\{\boldsymbol{h}_2, \ldots, \boldsymbol{h}_r\})\big\}$. Clearly, $|V| = q^{r-1}$. We show that every correctable erasure pattern is covered by some vector in V.

Let $S \subseteq [n]$ be correctable. Then $\mathrm{rank}\big(H(S)\big) = |S|$. Let $B \subseteq \{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_r\}$ be such that $B(S) \stackrel{\mathrm{def}}{=} \{\boldsymbol{b}(S) : \boldsymbol{b} \in B\}$ form a basis of $\mathbb{F}_q^{|S|}$. If $\boldsymbol{h}_1 \in B$, note that $\mathrm{span}\big(B(S) \setminus \boldsymbol{h}_1(S)\big)$ has dimension $|S|-1$, hence does not contain all weight-one $|S|$-tuples. Therefore, there exists $\boldsymbol{e} \in \mathbb{F}_q^{|S|}$, such that $\mathrm{wt}(\boldsymbol{e}) = 1$ and $\boldsymbol{e} = \alpha \boldsymbol{h}_1(S) + \boldsymbol{x}(S)$, where $\alpha \neq 0$ and $\boldsymbol{x} \in \mathrm{span}(B \setminus \boldsymbol{h}_1) \subseteq \mathrm{span}(\{\boldsymbol{h}_2, \ldots, \boldsymbol{h}_r\})$. Since $\mathrm{wt}(\alpha^{-1}\boldsymbol{e}) = 1$, we see S is covered by $(\boldsymbol{h}_1 + \alpha^{-1}\boldsymbol{x}) \in V$. Finally, if $\boldsymbol{h}_1 \notin B$, simply note that $V(S) = \mathrm{span}\big(B(S)\big) = \mathbb{F}_q^{|S|}$. ∎

Many techniques for bounding stopping redundancy can be applied to ML redundancy with little modification. We give just one such example in the following theorem, which is based on an idea from [5].

Theorem 6 *Let $\mathcal{C}$ be an $[n,k]_q$ code. Let $r = n-k$. Then for all $0 \leq p \leq 1$,*

$$\gamma(\mathcal{C}) \leq pq^r + \sum_{i=1}^{r} \binom{n}{i} (1-p)^{i(q-1)q^{r-i}}.$$

Proof: For some prescribed real value p, $0 \leq p \leq 1$, select each codeword in $\mathcal{C}^\perp$ with probability p, and let matrix H consist of all selected codewords as rows (with arbitrary ordering). At this point, the expected number of rows in H is

$$p \cdot |\mathcal{C}^\perp| = pq^r.$$

Not all correctable erasure patterns may be covered by H. Particularly, for any given $X \in [n]^i$, $i = 1, \ldots, r$, that is correctable, the probability that X is not covered by H is

$$(1-p)^{i(q-1)q^{r-i}}.$$

This is because X being correctable implies that $\mathcal{C}^\perp(X)$ (as a multiset) contains all q-ary i-tuples the same number of times. To cover all correctable erasure patterns, as a second step, for each $X \in [n]^i$, $i = 1, \ldots, r$, that is correctable but not yet covered by H, append to H a codeword from $\mathcal{C}^\perp$ that covers X, until no such X can be found. At this point, the expected number of rows in H is at most

$$pq^r + \sum_{i=1}^{r} \sum_{\substack{X \in [n]^i \\ X \text{correctable}}} (1-p)^{i(q-1)q^{r-i}}$$
$$\leq pq^r + \sum_{i=1}^{r} \binom{n}{i} (1-p)^{i(q-1)q^{r-i}}.$$

Therefore, there exists at least one realization of H with at most the above number of rows. ∎

We now take a different approach based on minimal codewords [11],[12]. A nonzero codeword is *minimal* if its support does not contain the support of another codeword of smaller weight.[3] The set of minimal codewords of code $\mathcal{C}$ is denoted by $\mathcal{M}(\mathcal{C})$. Certain properties [13],[12] of minimal codewords that are useful to us are summarized in the following lemma.

Lemma 7 *Let $\mathcal{C}$ be an $[n,k]_q$ code. Then*

1) *For all $\boldsymbol{v} \in \mathcal{C} \setminus \{\mathbf{0}\}$, there exist $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m \in \mathcal{M}(\mathcal{C})$, such that $\boldsymbol{v} = \sum_{i=1}^{m} \boldsymbol{c}_i$ and $\mathrm{supp}(\boldsymbol{c}_i) \subseteq \mathrm{supp}(\boldsymbol{v})$, $i = 1, \ldots, m$.*
2) *For all $\boldsymbol{c} \in \mathcal{M}(\mathcal{C})$, we have $\mathrm{wt}(\boldsymbol{c}) \leq n - k + 1$.*

We make a key observation that to achieve ML performance, it suffices to select parity checks from $\mathcal{M}(\mathcal{C}^\perp)$, which leads to the following upper bound on ML redundancy.

Theorem 8 *If $\mathcal{C}$ is a q-ary code, then*

$$\gamma(\mathcal{C}) \leq \frac{1}{q-1}\big|\mathcal{M}(\mathcal{C}^\perp)\big|.$$

Proof: Let $H_{\mathcal{M}}$ denote the matrix consisting of all codewords of $\mathcal{M}(\mathcal{C}^\perp)$ as rows. If $S \subseteq [n]$ is correctable, then by Lemma 1, it is covered by some $\boldsymbol{c} \in \mathcal{C}^\perp$. If $\boldsymbol{c} \in \mathcal{M}(\mathcal{C}^\perp)$, then $H_{\mathcal{M}}$ covers S. If $\boldsymbol{c} \notin \mathcal{M}(\mathcal{C}^\perp)$, then by Lemma 7, $\boldsymbol{c}$ can be written as a sum of codewords from $\mathcal{M}(\mathcal{C}^\perp)$, whose supports are contained in $\mathrm{supp}(\boldsymbol{c})$. Thus, one of these minimal codewords must cover S. Finally, note that just one codeword is needed from $\mathcal{M}(\mathcal{C}^\perp)$ for each support set. ∎

It is often hard to find the number of minimal codewords in $\mathcal{C}^\perp$. Lemma 7 tells us that it suffices to consider codewords of weights $k+1$ or less, leading to the following corollary. Note that the same result was also shown in [10, Theorem 3], albeit through a very different argument.

Corollary 9 *If $\mathcal{C}$ is an $[n,k]_q$ code, then*

$$\gamma(\mathcal{C}) \leq \frac{1}{q-1}\sum_{i=1}^{k+1} B_i,$$

where B_i is the number of weight-i codewords in $\mathcal{C}^\perp$.

Let us consider a few examples.

Example 1 Let $\mathcal{H}_m$ denote the binary Hamming code with redundancy m, and $\mathcal{S}_m$ denote its dual code, the $[2^m - 1, m, 2^{m-1}]_2$ Simplex code. Since all nonzero codewords of $\mathcal{S}_m$ have the same weight, all of them are minimal. Theorem 8 tells us that $\gamma(\mathcal{H}_m) \leq 2^m - 1$, which is trivially true and is about twice the true value of $\gamma(\mathcal{H}_m)$ as given in Theorem 4.

On the other hand, it is known [12] that the number of minimal codewords in $\mathcal{H}_m$ of weight w is

$$M_w = \begin{cases} \frac{1}{w!}\prod_{i=0}^{w-2}(2^m - 2^i) & \text{if } 3 \leq w \leq m+1, \\ 0 & \text{otherwise.} \end{cases}$$

So by Theorem 8, $\gamma(\mathcal{S}_m) \leq \big|\mathcal{M}(\mathcal{H}_m)\big| = O\big(2^{m^2}/m\big)$, which is much stronger than the upper bound of $2^{2^m - m - 2}$, as given by Theorem 3. □

[3]Some authors require in addition that the first nonzero coordinate of minimal codewords be one.

Example 2 Let $\mathcal{G}_{24}$ denote the $[24, 12, 8]_2$ self-dual Golay code. It is known [12] that $\mathcal{M}(\mathcal{G}_{24}) = \{\boldsymbol{c} \in \mathcal{G}_{24} \setminus \{\mathbf{0}\} : \mathrm{wt}(\boldsymbol{c}) \leq 12\}$. By Theorem 8 and the well-known weight distribution of $\mathcal{G}_{24}$ [14], we have $\gamma(\mathcal{G}_{24}) \leq 3335$. Note that Theorem 3 and Theorem 6 both perform better in this case, yielding $\gamma(\mathcal{G}_{24}) \leq 2048$ and $\gamma(\mathcal{G}_{24}) \leq 2435$, respectively. A greedy search shows that in fact $\gamma(\mathcal{G}_{24}) \leq 370$. □

Example 3 For MDS codes, minimal codewords and minimum-weight codewords are the same, because all nonzero codewords with weight not exceeding $r+1$ are minimum-weight. Also, ML redundancy and stopping redundancy become equivalent, since no correctable erasure pattern is of size larger than $r = d - 1$.

Let $\mathcal{C}$ be an $[n,k]_q$ MDS code. Then $\mathcal{C}^\perp$ is an $[n, n-k]_q$ MDS code. By Theorem 8, we have $\gamma(\mathcal{C}) \leq \binom{n}{k+1}$. Stronger upper and lower bounds on $\rho(\mathcal{C}) = \gamma(\mathcal{C})$ of MDS codes can be found in [4],[2],[5],[15]. □

Example 4 Let $\mathcal{C}$ be a random code whose $(n-k) \times n$ parity-check matrix consists of independent and equiprobable entries drawn from $\mathbb{F}_q$. The following result was shown in [12].

Theorem 10 ([12]) *Let $k = Rn$, where $R \in (0,1)$ is fixed. Then*

$$\lim_{n\to\infty} \frac{1}{n}\log_q \mathsf{E}\big[|\mathcal{M}(\mathcal{C})|\big] = \begin{cases} H_q(1-R) - (1-R) & \text{if } R > 1/q \\ R & \text{if } R \leq 1/q \end{cases}$$

where $H_q(\cdot)$ is the base-q entropy function.

Theorem 10 was proved by noting that the average number of minimal codewords of weight w is at least a constant fraction (> 0.288) of the average total number of weight-w codewords, for all $w \leq n-k+1$. This shows that Corollary 9 in general should give a bound that is not much larger than that given by Theorem 8.

To compare $\mathsf{E}\big[|\mathcal{M}(\mathcal{C})|\big]$ with the total number of codewords, note that $\mathsf{E}\big[|\mathcal{C}|\big] = q^k - q^{-r} + 1$. Therefore, if $R \leq 1/q$, almost all codewords in $\mathcal{C}$ are minimal; if $R > 1/q$, the number of minimal codewords as a fraction of $|\mathcal{C}|$ decreases to 0 exponentially fast as $n \to \infty$. Accordingly, Theorem 8 is likely to give us a non-trivial bound if $R < (q-1)/q$. This observation is corroborated by our earlier examples of Hamming, Simplex, and Golay codes. □

III. Codes with Designs

We first show that the ML redundancy of $\mathcal{S}_m$ is small — at most quadratic in the length of the code. This is contrasted with the very large number of available parity checks. It is also much lower than the $O\big(2^{m^2}/m\big)$ bound given by Theorem 8. Interestingly, the stopping redundancy of $\mathcal{S}_m$ is also very small. Indeed, it can be shown [6] that the stopping redundancy of $\mathcal{S}_m$ is equal to its redundancy, $2^m - m - 1$.

Theorem 11 *For all m,*

$$\gamma(\mathcal{S}_m) \leq \frac{n^2 - n}{6},$$

where $n = 2^m - 1$ is the block length of $\mathcal{S}_m$.

Proof: It is well known [14] that the number of weight-3 codewords in $\mathcal{H}_m$ is $\frac{1}{3}\binom{n}{2} = (n^2 - n)/6$. We will show

that any correctable erasure pattern for $\mathcal{S}_m$ is covered by a weight-3 codeword of $\mathcal{H}_m$. Let S be a correctable erasure pattern. Then there exists $\boldsymbol{c} \in \mathcal{H}_m$, such that $\mathrm{wt}\big(\boldsymbol{c}(S)\big) = 1$. If $\mathrm{wt}(\boldsymbol{c}) = 3$, then we are done. Otherwise, note that the supports of weight-3 codewords in $\mathcal{H}_m$ form a $(2, 3, n)$ Steiner system [14]. Hence, we can find a weight-3 codeword, $\boldsymbol{x}_1 \in \mathcal{H}_m$, that matches $\boldsymbol{c}$ at its nonzero position within S and at least one other nonzero bit. If $\mathrm{wt}\big(\boldsymbol{x}_1(S)\big) = 1$, then we are done. Otherwise, let $\boldsymbol{c}_1 = \boldsymbol{c} - \boldsymbol{x}_1$, and note that $\mathrm{wt}\big(\boldsymbol{c}_1(S)\big) = 1$ and $\mathrm{wt}(\boldsymbol{c}_1) < \mathrm{wt}(\boldsymbol{c})$. Now we can repeat the above procedure and find weight-3 codewords $\boldsymbol{x}_2, \boldsymbol{x}_3, \ldots \in \mathcal{H}_m$, until some $\boldsymbol{x}_i$ is found that covers S. This must happen since with each unsuccessful step the weight of the codeword that covers S (i.e. $\boldsymbol{c}_i = \boldsymbol{c}_{i-1} - \boldsymbol{x}_i$) is reduced by at least one. ∎

The above bound can be slightly improved by noting that a subset of the weight-3 codewords of $\mathcal{H}_m$ suffices as parity checks, giving the following result.

Theorem 12 *For all m,*

$$\gamma(\mathcal{S}_m) \le \frac{n^2 - 4n + 9}{6},$$

where $n = 2^m - 1$ is the block length of $\mathcal{S}_m$.

Proof: Let $\mathcal{A}_m$ denote the set of weight-3 codewords of $\mathcal{H}_m$. For $j = 1, \ldots, n$, define $\mathcal{A}_{m,j} \stackrel{\mathrm{def}}{=} \{\boldsymbol{x} \in \mathcal{A}_m : j \in \mathrm{supp}(\boldsymbol{x})\}$. Let $\mathcal{A}'_{m,j} \stackrel{\mathrm{def}}{=} \mathcal{A}_m \setminus \big(\mathcal{A}_{m,j} \setminus \{\boldsymbol{a}_{m,j}\}\big)$, where $\boldsymbol{a}_{m,j}$ is an arbitrary element in $\mathcal{A}_{m,j}$. We show that any correctable erasure pattern is covered by a vector in $\mathcal{A}'_{m,j}$ for all j.

Basically, we are saying that among all weight-3 codewords of $\mathcal{H}_m$ whose supports contain the j-th coordinate, we may discard all but one of them, and still cover all correct erasure patterns. Let $S \subseteq [n]$ be a correctable erasure pattern. Then there exists $\boldsymbol{c} \in \mathcal{H}_m$, such that $\mathrm{wt}\big(\boldsymbol{c}(S)\big) = 1$. Observe that if $j \notin \mathrm{supp}(\boldsymbol{c})$, then the recursive procedure in the previous proof can be followed through without using any vectors in $\mathcal{A}_{m,j}$. This is because at each step, we have multiple choices for $\boldsymbol{x}_i$ by varying the choice of the position outside of S, and only one of the resulting 3-sets possibly contains j.

We now show that either S is trivially covered, or we can find $\boldsymbol{c} \in \mathcal{H}_m$ such that $\mathrm{wt}\big(\boldsymbol{c}(S)\big) = 1$ and $j \notin \mathrm{supp}(\boldsymbol{c})$, so the above observation suffices to complete the proof. We need the following fact, which can be seen from the proof of Lemma 1:

$$\forall s \in S,\ \exists \boldsymbol{c} \in \mathcal{H}_m,\ \text{s.t. } \mathrm{supp}\big(\boldsymbol{c}(S)\big) = \{s\}. \qquad (2)$$

Now, consider two cases. If $j \in S$, then by (2), unless $S = \{j\}$ (in which case S is covered by $\boldsymbol{a}_{m,j}$), we can always choose $\boldsymbol{c} \in \mathcal{H}_m$ such that $\mathrm{wt}\big(\boldsymbol{c}(S)\big) = 1$ and $j \notin \mathrm{supp}(\boldsymbol{c})$. If $j \notin S$, let's further assume that $\mathrm{wt}\big(\boldsymbol{a}_{m,j}(S)\big) \neq 1$ (or S is covered by $\boldsymbol{a}_{m,j}$). Thus, $\mathrm{wt}\big(\boldsymbol{a}_{m,j}(S)\big) = 0$ or 2. If $\mathrm{wt}\big(\boldsymbol{a}_{m,j}(S)\big) = 0$, choose $\boldsymbol{v} \in \mathcal{H}_m$ such that $\mathrm{wt}\big(\boldsymbol{v}(S)\big) = 1$. If $\mathrm{wt}\big(\boldsymbol{a}_{m,j}(S)\big) = 2$, by (2) we can choose $\boldsymbol{v} \in \mathcal{H}_m$ such that $\mathrm{wt}\big(\boldsymbol{v}(S)\big) = 1$ and $\mathrm{supp}\big(\boldsymbol{v}(S)\big) \subset \mathrm{supp}(\boldsymbol{a}_{m,j})$. Now note that in either case, $\boldsymbol{v}$ and $\boldsymbol{v} + \boldsymbol{a}_{m,j}$ both cover S, but only one of them has j in its support. The one whose support does not contain j can then be chosen as $\boldsymbol{c}$.

Finally, since $\{\mathrm{supp}(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{A}_m\}$ is a $(2, 3, n)$ Steiner system, $|\mathcal{A}_{m,j}| = (n-1)/2$ for all j. Therefore, $|\mathcal{A}'_{m,j}| = |\mathcal{A}_m| - |\mathcal{A}_{m,j}| + 1 = n(n-1)/6 - (n-1)/2 + 1 = (n^2 - 4n + 9)/6$. ∎

The above bound is reasonably good, though not always sharp. For example, Theorem 12 shows that $\gamma(\mathcal{S}_3) \le 5$ and $\gamma(\mathcal{S}_4) \le 29$, while it can be verified through searching that $\gamma(\mathcal{S}_3) = 5$ and $\gamma(\mathcal{S}_4) \le 21$.

Next, consider Reed-Muller (RM) codes. Let $\mathcal{R}(l, m)$ denote the l-th order RM code with parameter m. The first order RM code, $\mathcal{R}(1, m)$, is closely related to the Simplex code $\mathcal{S}_m$. Whereas $\mathcal{S}_m$ is the dual code of the Hamming code $\mathcal{H}_m$, $\mathcal{R}(1, m)$ is the dual code of the corresponding extended Hamming code, which we denote by $\mathcal{H}_m^{\mathrm{ext}}$. Using ideas similar to those used to prove Theorem 11 and Theorem 12, we obtain the following bound on the ML redundancy of $\mathcal{R}(1, m)$.

Theorem 13 *For all m,*

$$\gamma\big(\mathcal{R}(1, m)\big) \le \frac{1}{4}\binom{n}{3} - \frac{n}{2} + 2.$$

where $n = 2^m$ is the block length of $\mathcal{R}(1, m)$.

Proof (Sketch): Let $\mathcal{B}_m$ denote the set of weight-4 codewords of $\mathcal{H}_m^{\mathrm{ext}}$. It is well known [14] that $\{\mathrm{supp}(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{B}_m\}$ is a $(3, 4, n)$ Steiner system. For each $D \in [n]^2$, define $\mathcal{B}_{m,D} \stackrel{\mathrm{def}}{=} \{\boldsymbol{x} \in \mathcal{B}_m : D \subset \mathrm{supp}(\boldsymbol{x})\}$, and let $\mathcal{B}'_{m,D} \stackrel{\mathrm{def}}{=} \mathcal{B}_m \setminus \big(\mathcal{B}_{m,D} \setminus \{\boldsymbol{b}_{m,D}\}\big)$, where $\boldsymbol{b}_{m,D}$ is an arbitrary element in $\mathcal{B}_{m,D}$.

Arguing as in the proofs of Theorem 11 and Theorem 12, one can show that any correctable erasure pattern is covered by a vector in $\mathcal{B}'_{m,D}$ for all D. Finally, note that $|\mathcal{B}'_{m,D}| = |\mathcal{B}_m| - |\mathcal{B}_{m,D}| + 1 = \frac{1}{4}\binom{n}{3} - \frac{n-2}{2} + 1$. ∎

The essence of Theorems 11–13 can be generalized to any length-n linear code whose dual code contains an $(n, t+1, t)$ covering design (by which we mean precisely that the supports of weight-$(t+1)$ codewords in the dual code form an $(n, t+1, t)$ covering design) for some t.

Theorem 14 *Let $\mathcal{C}$ be a linear code of length n over $\mathbb{F}_q$. If for some $t \ge q$, $\mathcal{C}^\perp$ contains an $(n, t+1, t)$ covering design, then*

$$\gamma(\mathcal{C}) \le \binom{n}{t} + \frac{1}{q-1}\sum_{i=1}^{t} B_i. \qquad (3)$$

In particular, if $\mathcal{C}^\perp$ contains a $(t, t+1, n)$ Steiner system, then

$$\gamma(\mathcal{C}) \le \frac{1}{t+1}\binom{n}{t} + \frac{1}{q-1}\sum_{i=1}^{t} B_i, \qquad (4)$$

where B_i is the number of weight-i codewords in $\mathcal{C}^\perp$.

Proof: Choose a subset of weight-$(t+1)$ codewords from $\mathcal{C}^\perp$ as parity checks, such that their supports form an $(n, t+1, t)$ covering design. This subset can be chosen to contain less than $\binom{n}{t}$ vectors (or, precisely $\frac{1}{t+1}\binom{n}{t}$ vectors, if $\mathcal{C}^\perp$ contains a $(t, t+1, n)$ Steiner system). Choose also a subset of codewords from $\mathcal{C}^\perp$ whose weights are less than $(t+1)$ (if there are any), such that exactly one codeword is chosen for each support set. The total number of vectors we have chosen from $\mathcal{C}^\perp$ is no greater than $\binom{n}{t} + \frac{1}{q-1}\sum_{i=1}^{t} B_i$ (or, $\frac{1}{t+1}\binom{n}{t} + \frac{1}{q-1}\sum_{i=1}^{t} B_i$, if $\mathcal{C}^\perp$ contains a $(t, t+1, n)$ Steiner system). Let this set of vectors (as parity checks) be denoted by $\mathcal{A}$. We show that any correctable erasure pattern for $\mathcal{C}$ is covered by a vector in $\mathcal{A}$.

If S is correctable, then there exists $\boldsymbol{v} \in \mathcal{C}^{\perp}$ such that $\mathrm{wt}\big(\boldsymbol{v}(S)\big) = 1$. Let $\boldsymbol{c} = \alpha\boldsymbol{v}$ where $\alpha \in \mathbb{F}_q$, $\alpha \neq 0$ is a constant of our choice. If there exists α such that $\boldsymbol{c} \in \mathcal{A}$, then we are done. Otherwise, we have $|\mathrm{supp}(\boldsymbol{c})| \geq t+1$ and $|\mathrm{supp}(\boldsymbol{c}) \setminus S| \geq t$. Let X be any t-subset of $\big(\mathrm{supp}(\boldsymbol{c}) \setminus S\big)$. Since $\mathcal{C}^{\perp}$ contains an $(n, t+1, t)$ covering design, there exists $\boldsymbol{x}_1 \in \mathcal{A}$ such that $\mathrm{wt}(\boldsymbol{x}_1) = t+1$ and $\mathrm{supp}(\boldsymbol{x}_1) \supset X$. If $\mathrm{wt}(\boldsymbol{x}_1(S)) = 1$, then we are done. Otherwise, note that by choosing α we can ensure that $\boldsymbol{c}$ matches $\boldsymbol{x}_1$ at no less than $\lceil t/(q-1) \rceil$ positions. Let $\boldsymbol{c}_1 = \alpha_1(\boldsymbol{c} - \boldsymbol{x}_1)$, where $\alpha_1 \in \mathbb{F}_q$, $\alpha_1 \neq 0$ can be chosen freely. Note that $\mathrm{wt}(\boldsymbol{c}_1(S)) = 1$. And since $t \geq q$, we also have

$$\mathrm{wt}(\boldsymbol{c}_1) \leq \mathrm{wt}(\boldsymbol{c}) - \lceil t/(q-1) \rceil + 1 < \mathrm{wt}(\boldsymbol{c}).$$

Now the above procedure can be repeated to find weight-$(t+1)$ codewords $\boldsymbol{x}_2, \boldsymbol{x}_3, \ldots \in \mathcal{A}$, and stopped if some $\boldsymbol{x}_i$ is found that covers S. Since at each unsuccessful step the weight of the parity-check vector that covers S (i.e. $\boldsymbol{c}_i = \alpha_i(\boldsymbol{c}_{i-1} - \boldsymbol{x}_i)$) is reduced by at least one, the process must terminate in a finite number of steps, at the end of which we will either have found some $\boldsymbol{x}_i \in \mathcal{A}$ that covers S, or we will be left with some $\boldsymbol{0} \neq \boldsymbol{c}_i = \alpha_i(\boldsymbol{c}_{i-1} - \boldsymbol{x}_i)$ that covers S and has weight no greater than t. In the latter case, note that by construction, there exists α_i such that $\boldsymbol{c}_i \in \mathcal{A}$. ■

The recursive procedure used in the above proof is slightly different from the one used in the proof of Theorem 11, but the two are similar in principle.

If the minimum distance of $\mathcal{C}^{\perp}$ is $t+1$, then the second term in (3) (and similarly in (4)) goes away, as in the case of Simplex and first order Reed-Muller codes.

As examples, let $\mathcal{G}_{11}$ and $\mathcal{G}_{12}$ denote the ternary and extended ternary Golay codes, respectively. It is well known [14] that the supports of weight-6 codewords in $\mathcal{G}_{12}$ and those of weight-5 codewords in $\mathcal{G}_{11}$ form $(5, 6, 12)$ and $(4, 5, 11)$ Steiner systems, respectively. Furthermore, note that $\mathcal{G}_{12}$ is self-dual. By (4), we immediately have $\gamma(\mathcal{G}_{11}^{\perp}) \leq 66$ and $\gamma(\mathcal{G}_{12}) \leq 132$.

IV. Concluding Remarks

We have defined and studied ML redundancy in the setting of erasure channels. A reasonable question to ask is how ML redundancy can be extended to other channels, such as the BSC or the AWGN channel. For such cases, it may be fruitful to consider how the choice of parity-check matrix affects the distribution of *pseudo-codewords* [16] (effectively assuming a linear programming decoder [17], whose performance is closely related to that of iterative decoding). One should be cautious, though, as there exist codes for which non-codeword pseudo-codewords persist even when all parity checks are deployed [17],[18]. For those codes, precise ML performance may never be matched using a linear programming decoder. Nevertheless, it could still be interesting to consider, for example, the smallest number of parity checks such that the volume of the fundamental polytope is minimized.

The recursive procedure used in proving Theorems 11 – 14 is reminiscent of the zero-neighbors algorithm [19]. It will be interesting to see if such intuition can be extended to the general case, so that instead of using all minimal codewords in the dual code (except those with duplicate supports), a (possibly much) smaller subset may suffice.

Acknowledgment

This work was supported in part by the Center for Magnetic Recording Research at UCSD, and by NSF Grant No. CCR-0514859.

References

[1] H. Pishro-Nik and F. Fekri, "On decoding of low-density parity-check codes over the binary erasure channel," *IEEE Trans. Inform. Theory*, vol. 50, no. 3, pp. 439–454, Mar. 2004.

[2] J. Han and P. H. Siegel, "Improved upper bounds on stopping redundancy," *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 90–104, Jan. 2007.

[3] H. D. L. Hollmann and L. M. G. M. Tolhuizen, "On parity check collections for iterative erasure decoding that correct all correctable erasure patterns of a given size," *IEEE Trans. Inform. Theory*, vol. 53, no. 2, pp. 823–828, Feb. 2007.

[4] M. Schwartz and A. Vardy, "On the stopping distance and the stopping redundancy of codes," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 922–932, Mar. 2006.

[5] J. Han, P. H. Siegel, and R. M. Roth, "Bounds on single-exclusion numbers and stopping redundancy of MDS codes," in *Proc. IEEE International Symposium on Information Theory*, Nice, France, June 2007, pp. 2941–2945.

[6] T. Etzion, "On the stopping redundancy of Reed-Muller codes," *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 4867–4879, Nov. 2006.

[7] O. Milenkovic, E. Soljanin, and P. Whiting, "Stopping and trapping sets in generalized covering arrays," in *Proc. 40th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2006, pp. 259–264.

[8] T. Hehn, S. Laendner, O. Milenkovic, and J. B. Huber, "The stopping redundancy hierarchy of cyclic codes," in *Proc. 44-th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, Sept. 2006, pp. 1271–1280.

[9] C. A. Kelley and D. Sridhara, "On the pseudocodeword weight and parity-check matrix redundancy of linear codes," in *Proc. IEEE Information Theory Workshop*, Lake Tahoe, California, Sept. 2007, pp. 1–6.

[10] J. H. Weber and K. A. S. Abdel-Ghaffar, "Results on parity-check matrices with optimal stopping and/or dead-end set enumerators," *IEEE Trans. Inform. Theory*, vol. 54, no. 3, pp. 1368–1374, Mar. 2008.

[11] T.-Y. Hwang, "Decoding linear block codes for minimizing word error rate," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 733–737, Nov. 1979.

[12] A. Ashikhmin and A. Barg, "Minimal vectors in linear codes," *IEEE Trans. Inform. Theory*, vol. 44, no. 5, pp. 2010–2017, Sept. 1998.

[13] J. L. Massey, "Minimal codewords and secret sharing," in *Proc. 6th Joint Swedish-Russian Workshop on Information Theory*, Rölle, Sweden, 1993, pp. 246–249.

[14] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland, 1978.

[15] J. Han, P. H. Siegel, and R. M. Roth, "Single exclusion number and stopping redundancy of MDS codes," submitted to *IEEE Trans. Inform. Theory*, 2007. [Online]. Available: http://arxiv.org/abs/0712.2857

[16] R. Koetter and P. Vontobel, "Graph covers and iterative decoding of finite-length codes," in *Proc. 3rd International Symposium on Turbo Codes and Related Topics*, Brest, France, Sept. 2003, pp. 75–82.

[17] J. Feldman, M. J. Wainwright, and D. R. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. Inform. Theory*, vol. 51, no. 3, pp. 954–972, Mar. 2005.

[18] F. Barahona and M. Grötschel, "On the cycle polytope of a binary matroid," *J. Combin. Theory Ser. B*, vol. 40, pp. 40–62, 1986.

[19] L. B. Levitin and C. R. P. Hartmann, "A new approach to the general minimum distance decoding problem: The zero-neighbors algorithm," *IEEE Trans. Inform. Theory*, vol. 31, no. 3, pp. 378–384, May 1985.