The investigation was carried out by formulating the problem in a discrete form. This leads to the representation of a quadriphase code by discrete Fourier transform and inverse discrete Fourier transform. The coefficients of a mismatched filter are then calculated from (7) by exploiting the discrete Fourier transform representation of the code. The mismatched filter designed in this manner eliminates the sidelobes as we have shown in Figs. 1, 2, and 6.

A mismatched filter has three drawbacks when compared to the matched filter. The first drawback is that a mismatched filter has infinitely many coefficients. However, we have explained and also shown in Fig. 3 that the values of these coefficients go to zero rapidly. Hence, this may not be problematic in practice. The second drawback is that the discrete Fourier transform of the code might have zeros in the frequency domain and we have shown that this is improbable. The third problem is that its output $\mathbf{SNR}$ is less than the output $\mathbf{SNR}$ of the respective matched filter. However, one can choose codes with minimum $\mathbf{SNR}$ losses. We have presented the noise enhancement factors for the best codes in Table I. These codes have been selected from 1.466×10^{12} number of investigated codes based on their k_{vs} values. The quadriphase codes, which have smaller values of k_{vs} than the respective binary-phase codes, were selected. For code lengths of $7, 10, 9, 15, 16, 17, 18,$ and 19, the optimal codes turned out to be quadriphase codes. The other optimal codes are binary-phase codes.

The 13- and 15-element codes have 0.2136 and 0.3012 noise enhancement factors, respectively. For these codes we have shown that the shape of the mismatched filter resembles the matched filter. In most of the other codes, the noise enhancement factor is greater than 0.4. One should evaluate the losses in $\mathbf{SNR}$ against the advantages of eliminating sidelobes. Finally, by carrying out a random search we have shown that a randomly selected long code will most likely have large noise enhancement factor.

REFERENCES

[1] R. H. Barker, "Group synchronizing of binary digital systems," in *Communications Theory*. London, U.K.: Butterworth, 1953, pp. 273–287.

[2] R. J. Turyn, "Sequences with small correlation," in *Error Correcting Codes*, H. B. Mann, Ed. New York: Wiley, 1968, pp. 195–228.

[3] J. Lindner, "Binary sequences up to length 40 with possible autocorrelation function," *Electron. Lett.*, vol. 11, pp. 507–507, 1975.

[4] Coxson and Russo, "Efficient exhaustive search for optimal-peak-sidelobe binary sequences," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 1, pp. 302–308, Jan. 2005.

[5] E. L. Key, E. N. Fowle, and R. D. Haggarty, A Method of Sidelobe Suppression in Phase Coded Pulse Compression Systems, MIT, Lincoln Lab., Lexington, MA, Tech Rep. 209, Nov. 1959.

[6] H. Rohling and W. Plagge, "Mismatched-filter design for periodical binary phased signals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 6, pp. 890–896, Nov. 1989.

[7] M. S. Lehtinen, B. Damtie, and T. Nygrén, "Optimal binary phase codes and sidelobe-free decoding filters with application to incoherent scatter radar," *Ann. Geophysicae*, vol. 22, pp. 1623–1632, 2004.

[8] M. S. Lehtinen and I. Häggström, "A new modulation principle for incoherent scatter measurements," *Radio Sci.*, vol. 22, pp. 625–634, 1987.

[9] J. Rupprecht and M. Rupf, "On the search for good aperiodic binary invertible sequences," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1604–1612, Sep. 1996.

[10] J. Rupprecht, "Maximum-Likelihood Estimation of Multipath Channels," Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich, Swwitzerland, 1989.

[11] S. W. Golomb and R. A. Scholtz, "Generalized barker sequences," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 4, pp. 533–537, Oct. 1965.

[12] R. J. Turyn, "Four-phase barker codes," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 3, pp. 366–371, May 1974.

[13] J. W. Taylor, Jr. and H. J. Blinchikoff, "Quadriphase code-a radar pulse compression signal with unique characteristics," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 2, pp. 156–170, Mar. 1988.

[14] W. H. Mow, "Best quadriphase codes up to length 24," *Electron. Lett.*, vol. 29, pp. 923–925, 1993.

[15] H. D. Lüke, "Mismatched filtering of periodic quadriphase and 8-phase sequences," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1061–1063, Jul. 2003.

[16] R. Courant and D. Hilbert, *Methoden der mathematischen Physik*. Berlin, Germany: Springer-Verlag, 1968.

Improved Probabilistic Bounds on Stopping Redundancy

Junsheng Han, *Student Member, IEEE*, Paul H. Siegel, *Fellow, IEEE*, and Alexander Vardy, *Fellow, IEEE*

***Abstract*—For a linear code $\mathbb{C}$, the *stopping redundancy* of $\mathbb{C}$ is defined as the minimum number of check nodes in a Tanner graph T for $\mathbb{C}$ such that the size of the smallest stopping set in T is equal to the minimum distance of $\mathbb{C}$. Han and Siegel recently proved an upper bound on the stopping redundancy of general linear codes, using probabilistic analysis. For most code parameters, this bound is the best currently known. In this correspondence, we present several improvements upon this bound.**

***Index Terms*—Binary erasure channel, iterative decoding, linear codes, stopping redundancy, stopping sets.**

I. INTRODUCTION

The stopping redundancy of a binary linear code $\mathbb{C}$ characterizes the complexity, measured in terms of the minimum number of check nodes, of a Tanner graph for $\mathbb{C}$ such that iterative decoding of this graph on the binary erasure channel (BEC) achieves performance comparable (up to a constant factor) to that of maximum-likelihood decoding. It is widely believed [16], [19] that stopping redundancy and related concepts are of relevance for channels other than the BEC as well.

Specifically, let $\mathbb{C}$ be a linear code of length n, dimension k, and minimum distance d, and let $H = [h_{ij}]$ be a $t \times n$ parity-check matrix[1] for $\mathbb{C}$. The corresponding Tanner graph $T(H)$ for $\mathbb{C}$ is a bipartite graph with n variable nodes and t check nodes such that the ith check node is adjacent to the jth variable node iff $h_{ij} \neq 0$. A *stopping set* $\mathcal{S}$ is

Manuscript received July 27, 2007; revised December 13, 2007. This work was supported in part by the National Science Foundation.

J. Han and P. H. Siegel are with the Center for Magnetic Recording Research, University of California San Diego, La Jolla, CA 92093 USA (e-mail: han@cts.ucsd.edu; psiegel@ucsd.edu).

A. Vardy is with the Departments of Electrical Engineering and Computer Science, University of California San Diego, La Jolla, CA 92093-0407 USA (e-mail: vardy@kilimanjaro.ucsd.edu).

Communicated by T. Etzion, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2008.917624

[1]Throughout this correspondence, a *parity-check matrix for* $\mathbb{C}$ should be understood as any matrix H whose rows span the dual code $\mathbb{C}^{\perp}$. Thus t can be (and often will be) strictly greater than $n - k$.

a subset of the variable nodes in $T(H)$ such that all the check nodes that are neighbors of $\mathcal{S}$ are connected to $\mathcal{S}$ at least twice. Equivalently, a stopping set $\mathcal{S}$ is a set of columns of the parity-check matrix such that the corresponding column submatrix of H does not contain a row of weight one. It is well known [3], [19] that iterative decoding on the BEC succeeds iff the set of erased positions does not contain a stopping set. The size of the smallest stopping set, known as the *stopping distance* $s(H)$, is thus a limit on the number of erasures that iterative decoding on the Tanner graph $T(H)$ can guarantee to correct.

Note that the stopping distance is not a property of the code $\mathbb{C}$ itself, but rather of the specific choice of a parity-check matrix H for $\mathbb{C}$. Moreover, it is known [19] that $s(H) \leqslant d$ for all possible choices of H, and it is always possible to find a parity-check matrix for $\mathbb{C}$ such that this bound holds with equality. This motivates the following definition.

Definition 1: Let $\mathbb{C}$ be a linear code with minimum Hamming distance d. Then the ***stopping redundancy*** of $\mathbb{C}$ is defined as the the smallest integer $\rho(\mathbb{C})$ such that there exists a parity-check matrix H for $\mathbb{C}$ with $\rho(\mathbb{C})$ rows and $s(H) = d$.

Stopping redundancy was introduced by Schwartz and Vardy in [18], [19]. It was subsequently studied in a number of recent papers [1], [4], [6]–[13], [17] and [20]. Related concepts, such as stopping redundancy hierarchy, trapping redundancy, and generic erasure-correcting matrices were investigated in [8]–[13], [16], and [20]. Existing results on stopping redundancy are of two types: bounds on the stopping redundancy of specific families of codes (e.g., cyclic codes [8], MDS codes [6], [7], [19], Reed–Muller codes [4], [19], and Hamming codes [4], [20]) as well as bounds on the stopping redundancy of *general (binary) linear codes* [6], [10]–[13], [18], [19]. It is the latter type of results that we are concerned with here. Thus let us describe the known bounds in some detail.

Let $\mathbb{C}$ be an (n, k, d) binary linear code, and let $r = n - k$. The first upper bound on the stopping redundancy of $\mathbb{C}$, established by Schwartz and Vardy [18], [19, Theorem 4], is given by

$$\rho(\mathbb{C}) \leqslant \binom{r}{1} + \binom{r}{2} + \cdots + \binom{r}{d-2} \tag{1}$$

provided $d \geqslant 3$. This bound was improved (at least, for odd d) by Han and Siegel [6, Theorem 2], yielding

$$\rho(\mathbb{C}) \leqslant \binom{r}{1} + \binom{r}{3} + \binom{r}{5} + \cdots + \binom{r}{2t-1} \tag{2}$$

where $t = \lfloor d/2 \rfloor$. Hollmann and Tolhuizen [10]–[13] studied a more difficult problem of constructing "generic erasure-correcting sets." However, some of their results can be interpreted as bounds on the stopping redundancy hierarchy (cf. [8], [9]) of general linear codes, and then specialized to bounds on $\rho(\mathbb{C})$. In particular, Hollmann and Tolhuizen prove in [12, Theorem 5.2] and [13, Theorem 4.1] that

$$\rho(\mathbb{C}) \leqslant \binom{r-1}{0} + \binom{r-1}{1} + \cdots + \binom{r-1}{d-2} \tag{3}$$

which improves upon (1) and (2). They also show in [10, Theorem 1] that if all the codewords of $\mathbb{C}$ have even weight, then

$$\rho(\mathbb{C}) \leqslant 2\left(\binom{r-2}{0} + \binom{r-2}{1} + \cdots + \binom{r-2}{d-3}\right). \tag{4}$$

The bounds in (1)–(4) are all based on the same general set of ideas, and are constructive in a certain well-defined sense (cf. [12], [13]). In contrast, Hollmann and Tolhuizen employed a nonconstructive random-coding argument in [11, Theorem 4.2] and [13, Theorem 3.2] to show that

$$\rho(\mathbb{C}) \leqslant \frac{\log_2((2^r-1)(2^r-2)(2^r-2^2)\cdots(2^r-2^{d-2}))}{(d-1)-\log_2(2^{d-1}-(d-1))}. \tag{5}$$

As pointed out by Ludo Tolhuizen, another probabilistic bound on stopping redundancy, namely

$$\rho(\mathbb{C}) \leqslant \frac{(r-1)(d-2)-\log_2((d-2)!)}{(d-2)-\log_2(2^{d-2}-1)} + 1 \tag{6}$$

follows by combining Proposition 6.2, Lemma 6.5, and Theorem 6.14 of [12]. An entirely different probabilistic argument was used by Han and Siegel in [6, Theorem 3] to establish the following bound:

$$\rho(\mathbb{C}) \leqslant \min\left\{t \in \mathbb{N} : \mathcal{E}_{n,d}(t) < 1\right\} + (r - d + 1) \tag{7}$$

where

$$\mathcal{E}_{n,d}(t) \stackrel{\text{def}}{=} \sum_{i=1}^{d-1} \binom{n}{i} \left(1 - \frac{i}{2^i}\right)^t. \tag{8}$$

Although the various bounds in (1)–(7) are difficult to compare analytically, the following claims seem to be justified.

- For most code parameters, the probabilistic bounds (5)–(7) are much better than the constructive bounds (1)–(4).
- Among the probabilistic bounds, the Han–Siegel bound of [6, Theorem 3] and (7) is by far the strongest.

Using the tables of [5], we have verified that these claims hold for *all* possible code parameters n, k, and d with $d > 4$ and $n \leqslant 256$. A representative picture is shown in Table I, where we computed the bounds (1)–(7) for the $(24, 12, 8)$ Golay code and for the $(48, 24, 12)$ quadratic-residue code. Thus the Han–Siegel bound (7) appears to be the best currently known bound on the stopping redundancy of general linear codes.

TABLE I
UPPER BOUNDS ON THE STOPPING REDUNDANCY OF THE $(24, 12, 8)$ AND THE $(48, 24, 12)$ CODES

	$(24, 12, 8)$ Golay	$(48, 24, 12)$ QR
Schwartz-Vardy (1)	2,509	4,540,385
Han-Siegel (2)	1,816	4,194,304
Hollmann-Tolhuizen (3)	1,486	2,842,226
Hollmann-Tolhuizen (4)	1,276	2,195,580
Hollmann-Tolhuizen (6)	2,488	147,712
Hollmann-Tolhuizen (5)	1,034	33,978
Han-Siegel (7)	232	4,440

In this correspondence, we present several improvements upon the Han–Siegel bound (7) using a number of ideas, most of which are based upon a more careful probabilistic analysis.

II. PROBABILISTIC BOUNDS ON STOPPING REDUNDANCY

Given a linear code $\mathbb{C}$, Han and Siegel [6, Theorem 3] construct a parity-check matrix H for $\mathbb{C}$ by drawing codewords independently and uniformly at random from the dual code $\mathbb{C}^\perp$. Our first observation is this: such random choice is efficient at first, but becomes less and less efficient as successive rows of H are drawn from $\mathbb{C}^\perp$. At some point, deterministic selection becomes superior. This leads to the following result.

Theorem 1: Let $\mathbb{C}$ be an (n,k,d) binary linear code. Let $\Delta(\mathbb{C})$ denote the *deficiency* of $\mathbb{C}$, that is $\Delta(\mathbb{C}) = n-k+1-d$. Let $\mathcal{E}_{n,d}(t)$ be the expectation function defined in (8). Then

$$\rho(\mathbb{C}) \leqslant \min_{t\in\mathbb{N}} \left\{ t + \lfloor \mathcal{E}_{n,d}(t) \rfloor \right\} + \Delta(\mathbb{C}). \tag{9}$$

Proof: We first need some notation. Let $\binom{[n]}{i}$ denote the set of all subsets of $\{1,2,\ldots,n\}$ of size i, and let

$$\mathfrak{U} \stackrel{\text{def}}{=} \bigcup_{i=1}^{d-1} \binom{[n]}{i}. \tag{10}$$

Following [6], [19] we will refer to elements of $\binom{[n]}{i}$ as *i-sets*. Given an $m \times n$ binary matrix H and an i-set $\mathcal{S}$, we say that H *covers* $\mathcal{S}$ if the $m \times i$ column submatrix of H, consisting of those columns that are indexed by the elements of $\mathcal{S}$, contains a row of weight one. Clearly, the stopping sets with respect to H are precisely those sets $\mathcal{S}$ that H does not cover; hence $s(H) \geqslant d$ if and only if H covers $\mathcal{S}$ for all $\mathcal{S} \in \mathfrak{U}$.

Now let H_t be a $t \times n$ matrix whose rows are drawn from $\mathbb{C}^\perp$ independently and uniformly at random. If $\mathcal{S} \in \mathfrak{U}$ is a fixed i-set and $\boldsymbol{h}$ is a fixed row of H_t, then the probability that $\boldsymbol{h}$ covers $\mathcal{S}$ is $i/2^i$. This is so because $\mathbb{C}^\perp$ is an orthogonal array of strength $d-1$ (cf. [15, p. 139]), which means precisely that for all $\mathcal{S} \in \mathfrak{U}$, a codeword drawn at random from $\mathbb{C}^\perp$ is equally likely to contain each of the 2^i possible vectors on the i positions indexed by $\mathcal{S}$. Let X_t denote the number of sets $\mathcal{S} \in \mathfrak{U}$ that are not covered by H_t. Then X_t is a random variable, and the expected value of X_t is given by

$$\begin{aligned} \mathsf{E}[X_t] &= \sum_{\mathcal{S}\in\mathfrak{U}} \Pr\{\mathcal{S} \text{ not covered by } H_t\} \\ &= \sum_{i=1}^{d-1} \binom{n}{i} \left(1 - \frac{i}{2^i}\right)^t. \end{aligned}$$

The right-hand side of the above expression is $\mathcal{E}_{n,d}(t)$, by definition. It follows that there exists a realization H of H_t that covers all but at most $\lfloor \mathcal{E}_{n,d}(t) \rfloor$ sets in $\mathfrak{U}$. For each uncovered set $\mathcal{S}$, there is a codeword of $\mathbb{C}^\perp$ that covers $\mathcal{S}$ (again, since $\mathbb{C}^\perp$ is an orthogonal array). Thus we can adjoin $\lfloor \mathcal{E}_{n,d}(t) \rfloor$ rows to H to create a matrix H' with $s(H') = d$. It is possible that $\operatorname{rank}(H') < n-k$. However, $\operatorname{rank}(H')$ is clearly at least $d-1$. Hence, by adjoining at most $\Delta(\mathbb{C}) = (n-k)-(d-1)$ rows to H', we finally obtain a parity-check matrix H'' for $\mathbb{C}$ with at most $t + \lfloor \mathcal{E}_{n,d}(t) \rfloor + \Delta(\mathbb{C})$ rows and $s(H'') = d$. □

Since the minimization in (9) contains the Han-Siegel bound of (7) as a special case, Theorem 1 is at least as strong as (7). In fact, it is often substantially stronger (cf. Table II).

The bound of Theorem 1 involves solving a minimization problem; a closed-form expression would be desirable. This is addressed in the following corollary, which is similar in spirit to [6, Corollary 4] and improves upon it. The corollary provides an approximate closed-form solution for the minimization problem of (9), which becomes exact asymptotically.

Corollary 2: Let $\mathbb{C}$ be an (n,k,d) binary linear code. Let $\Delta(\mathbb{C})$ be the deficiency of $\mathbb{C}$, as before, and define

$$\begin{aligned} C &\stackrel{\text{def}}{=} \sum_{i=1}^{d-1} \binom{n}{i} \simeq 2^{nH_2(\frac{d}{n})} \\ D &\stackrel{\text{def}}{=} -\ln\left(1 - \frac{d-1}{2^{d-1}}\right) \simeq \frac{d-1}{2^{d-1}} \end{aligned}$$

TABLE II
IMPROVED BOUNDS ON THE STOPPING REDUNDANCY OF THE $(24,12,8)$ AND THE $(48,24,12)$ CODES

	$(24,12,8)$ Golay	$(48,24,12)$ QR
Han-Siegel (7)	232	4,440
Theorem 1	198	3,655
Corollary 2	213	3,738
Theorem 3	194	3,655
Theorem 4	187	3,577
Proposition 5 Theorem 7	182	3,564

where $H_2(x) \stackrel{\text{def}}{=} -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function. Then

$$\rho(\mathbb{C}) \leqslant \left\lceil \frac{\ln C + \ln D + 1}{D} \right\rceil + \Delta(\mathbb{C}). \tag{11}$$

Proof: In order to solve the minimization problem in (9), albeit approximately, we upper-bound $\mathcal{E}_{n,d}(t)$ as follows:

$$\mathcal{E}_{n,d}(t) \leqslant \sum_{i=1}^{d-1} \binom{n}{i} \left(1 - \frac{d-1}{2^{d-1}}\right)^t = Ce^{-Dt}$$

where the inequality follows from the fact that $i/2^i$ is nonincreasing in i. Let $f(t) = t + Ce^{-Dt}$. Then $f(t)$, regarded as a function from $\mathbb{R}$ to $\mathbb{R}$, is convex and attains its global minimum at $t_0 = (\ln C + \ln D)/D$. Clearly, t_0 is always positive and $f(t_0) = (\ln C + \ln D + 1)/D$. The corollary now follows from Theorem 1. □

Next, we observe that the random choice method used by Han and Siegel in [6, Theorem 3] and in Theorem 1 is not optimal. It would be better to select the rows of H_t from $\mathbb{C}^\perp \setminus \{\mathbf{0}\}$, rather than $\mathbb{C}^\perp$, without replacement, rather than with replacement. This leads to the following bound.

Theorem 3: Let $\mathbb{C}$ be an (n,k,d) binary linear code, let $\Delta(\mathbb{C})$ be the deficiency of $\mathbb{C}$, and let $r = n-k$. Then

$$\rho(\mathbb{C}) \leqslant \min_{t\in\mathbb{N}} \left\{ t + \lfloor \mathcal{F}_{n,d,k}(t) \rfloor \right\} + \Delta(\mathbb{C}) \tag{12}$$

where

$$\mathcal{F}_{n,d,k}(t) \stackrel{\text{def}}{=} \sum_{i=1}^{d-1} \binom{n}{i} \prod_{j=1}^{t} \left(1 - \frac{i2^{r-i}}{2^r - j}\right). \tag{13}$$

Proof: We construct a parity-check matrix for $\mathbb{C}$ in the same way as in Theorem 1, except that the initial matrix H_t is selected differently. Specifically, the rows $\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_t$ of H_t are selected as follows. Suppose that we have already chosen the first $j-1$ rows $\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_{j-1}$, then the jth row $\boldsymbol{h}_j$ is selected uniformly at random from

$$\mathbb{C}^\perp \setminus \{\mathbf{0}, \boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_{j-1}\}. \tag{14}$$

Now let $\mathcal{S} \in \mathfrak{U}$ be a fixed i-set that is not covered by any of $\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_{j-1}$. What is the probability that the jth row covers $\mathcal{S}$? The total number of possible (equally likely) choices for $\boldsymbol{h}_j$ is $2^r - j$. Of these, precisely $i2^{r-i}$ cover $\mathcal{S}$. This is so because there are exactly $i2^{r-i}$ codewords in $\mathbb{C}^\perp$ that cover $\mathcal{S}$ (again, since $\mathbb{C}^\perp$ is an orthogonal array), and none of these is among $\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_{j-1}$ by assumption.

Hence, the probability that $\boldsymbol{h}_j$ covers $\mathcal{S}$ is $i2^{r-i}/(2^r - j)$. Thus, for any fixed i-set $\mathcal{S} \in \mathfrak{U}$, the probability that H_t does not cover $\mathcal{S}$ is

$$\Pr\{\mathcal{S} \text{ is not covered by } H_t\} = \prod_{j=1}^{t} \left(1 - \frac{i2^{r-i}}{2^r - j}\right).$$

Thus, $\mathcal{F}_{n,d,k}(t)$ in (13) is the expected number of sets $\mathcal{S} \in \mathfrak{U}$ that are not covered by H_t, and the theorem follows. □

Since (13) is extremely close to (8), it is somewhat surprising that Theorem 3 yields any improvements upon Theorem 1. But it does, at least for small code parameters (cf. Table II).

We now improve the construction of a parity-check matrix for $\mathbb{C}$ described in Theorem 3 in yet another respect. Let H_0 be a fixed $t \times n$ matrix whose rows $\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_t$ are nonzero codewords of $\mathbb{C}^\perp$, and let $\mathfrak{U}_0$ denote the subset of the set $\mathfrak{U}$ in (10) consisting of those sets that are not covered by H_0. Let $X_0 = |\mathfrak{U}_0|$. Suppose we adjoin another row $\boldsymbol{h}_{t+1}$ to H_0, selected uniformly at random from $\mathbb{C}^\perp \setminus \{\boldsymbol{0}, \boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_t\}$ as in (14), and let H_1' denote the resulting random matrix. Let X_1' be the number of sets $\mathcal{S} \in \mathfrak{U}$ not covered by H_1'. Then, arguing as in the proof of Theorem 3, we find that

$$\begin{aligned}\mathsf{E}[X_1'] &= \sum_{\mathcal{S} \in \mathfrak{U}_0} \Pr\{\mathcal{S} \text{ is not covered by } \boldsymbol{h}_{t+1}\} \\ &\leqslant X_0 \left(1 - \frac{(d-1)2^{r-d+1}}{2^r - (t+1)}\right).\end{aligned}$$

It follows that there exists a $(t+1) \times n$ realization H_1 of H_1' that covers all but at most

$$X_1 = \left\lfloor X_0 \left(1 - \frac{(d-1)2^{r-d+1}}{2^r - (t+1)}\right)\right\rfloor \tag{15}$$

sets in $\mathfrak{U}$. The process can be now iterated. That is, given H_1, we can construct a $(t+2) \times n$ matrix H_2 that covers all but at most

$$X_2 = \left\lfloor X_1 \left(1 - \frac{(d-1)2^{r-d+1}}{2^r - (t+2)}\right)\right\rfloor \tag{16}$$

sets in $\mathfrak{U}$. And so on. To formalize this process, let us define for all $j = 1, 2, \ldots$, the function $P_j : \mathbb{N} \to \mathbb{N}$ as follows:

$$P_j(k) \stackrel{\text{def}}{=} \left\lfloor k \left(1 - \frac{(d-1)2^{r-d+1}}{2^r - (t+j)}\right)\right\rfloor, \quad \text{for all } k \in \mathbb{N}$$

where the parameters d, r, and t are regarded as constant. Then, after i iterations of the foregoing procedure, we will construct a $(t+i) \times n$ matrix H_i that covers all but at most

$$Q_i(X_0) \stackrel{\text{def}}{=} P_i\Big(P_{i-1}\Big(\cdots P_2\Big(P_1(X_0)\Big)\Big)\Big) \tag{17}$$

sets in $\mathfrak{U}$. If $Q_i(X_0) = 0$, we are done. This establishes the following upper bound on stopping redundancy.

Theorem 4: Let $\mathbb{C}$ be an (n, k, d) binary linear code, with deficiency $\Delta(\mathbb{C})$. Then the stopping redundancy $\rho(\mathbb{C})$ is at most

$$\min_{t \in \mathbb{N}} \Big\{t + \min\{i \in \mathbb{N} : Q_i(\lfloor \mathcal{F}_{n,d,k}(t) \rfloor) = 0\}\Big\} + \Delta(\mathbb{C}) \tag{18}$$

where $\mathcal{F}_{n,d,k}(t)$ is the expectation function defined in (13) while $Q_i(\cdot)$ is the function defined in (17).

Although the definition of $Q_i(\cdot)$ in (17) appears to be rather involved, we observe that the minimization over i in (18) is, in fact, very easy to compute: all it takes is a single-line while loop. The bound of Theorem 4 clearly includes Theorem 3 as a special case, and improves upon it (cf. Table II).

Finally, we would like to get rid of the small, but annoying, deficiency term $\Delta(\mathbb{C})$ in Theorems 1–4. To this end, the following simple observation often suffices. A code $\mathbb{C} \subseteq \mathbb{F}_2^n$ is said to be *maximal* if it is not possible to adjoin any vector in $\mathbb{F}_2^n$ to $\mathbb{C}$ without decreasing its minimum distance.

Proposition 5: Let $\mathbb{C}$ be an (n, k, d) binary linear code. Let H be any matrix with $s(H) = d$ whose rows are codewords of $\mathbb{C}^\perp$. If $\mathbb{C}$ is maximal, then $\operatorname{rank}(H) = n - k$.

Proof: Assume to the contrary that $\operatorname{rank}(H) < n - k$. Then H is a parity-check matrix for a proper supercode $\mathbb{C}'$ of $\mathbb{C}$. Since $s(H) = d$, the minimum distance of $\mathbb{C}'$ is also d, which contradicts the fact that $\mathbb{C}$ is maximal. □

Proposition 5 implies that if $\mathbb{C}$ is maximal, we can drop the deficiency term $\Delta(\mathbb{C})$ in Theorems 1–4. We can also get rid of this term using a more elaborate probabilistic argument. It is intuitively clear that if we draw sufficiently many codewords uniformly at random from the dual of an (n, k, d) code $\mathbb{C}$, the resulting matrix is likely to have rank close to $r = n - k$. This observation is made formal in the following lemma.

Lemma 6: Let $\mathbb{C}$ be an (n, k, d) binary linear code, and let H_t be a $t \times n$ matrix whose rows are drawn uniformly at random (either with or without replacement) from the dual code $\mathbb{C}^\perp$. Let $r = n - k$, and define the random variable $Y_t = r - \operatorname{rank}(H_t)$. Then for all $t \geqslant r$, we have

$$\mathsf{E}[Y_t] \leqslant \frac{1}{2^{t-r}} \left(1 + \frac{2/3}{2^{t-r+1} - 1}\right). \tag{19}$$

Proof Sketch: First assume that the rows of H_t are drawn from $\mathbb{C}^\perp$ uniformly at random *with* replacement (as in Theorem 1). It is known (see [14], for example) that

$$\Pr\{\operatorname{rank}(H_t) = j\} = \frac{1}{2^{(r-j)(t-j)}} \prod_{i=0}^{j-1} \frac{(1 - 2^{i-r})(1 - 2^{i-t})}{1 - 2^{i-j}}$$

for all $j = 1, 2, \ldots, r$. Taking the expectation with respect to this distribution (along with some technical details that we omit) produces the upper bound in (19). Finally, it can be shown that if the rows of H_t are drawn uniformly at random from $\mathbb{C}^\perp \setminus \{\boldsymbol{0}\}$ *without* replacement (as in Theorem 3), this can only reduce the expected value of Y_t. □

In order to combine Lemma 6 with our earlier results, let us define the function

$$\mathcal{G}_{n,d,k}(t) \stackrel{\text{def}}{=} \mathcal{F}_{n,d,k}(t) + \frac{1}{2^{t-r}} \left(1 + \frac{2/3}{2^{t-r+1} - 1}\right) \tag{20}$$

where $r = n - k$ and $\mathcal{F}_{n,d,k}(t)$ is as defined in (13). We can now modify the proofs of Theorem 3 and Theorem 4 accordingly, thereby establishing the following result.

Theorem 7: Let $\mathbb{C}$ be an (n, k, d) binary linear code, and let $r = n - k$. Then the stopping redundancy of $\mathbb{C}$ is bounded by

$$\rho(\mathbb{C}) \leqslant \min_{t \geq r} \Big\{t + \lfloor \mathcal{G}_{n,d,k}(t) \rfloor\Big\}. \tag{21}$$

Moreover, if $(r-1)(d-1) \leqslant 2^{d-1}$ then

$$\rho(\mathbb{C}) \leqslant \min_{t \geq r} \Big\{t + \min\{i \in \mathbb{N} : Q_i(\lfloor \mathcal{G}_{n,d,k}(t) \rfloor) = 0\}\Big\} \tag{22}$$

where the functions $\mathcal{G}_{n,d,k}(t)$ and $Q_i(\cdot)$ are as defined in (20) and (17), respectively.

Proof: Use the same argument as in the proof of Theorem 3, but with respect to the random variable $Z_t = X_t + Y_t$, which is the sum of the number of sets $\mathcal{S} \in \mathfrak{U}$ not covered by H_t and its rank deficiency $r - \mathrm{rank}(H_t)$. Further, note that the condition $(r-1)(d-1) \leqslant 2^{d-1}$ is sufficient for the argument in (15), (16), and (17) to be applicable in this case as well. □

It is not immediately apparent that (21) produces a better bound on stopping redundancy than Theorem 3. However, we can show that this is *always* so, except for a few trivial cases. In fact, comparing (13) and (19), we see that the second term in (20) decreases with t exponentially faster than the first term. Thus, unless the minimum in (21) and/or (22) is achieved for t very close to r (in which case $\rho(\mathbb{C})$ must be close to r as well), the second term in (20) has essentially no effect on the minimization—this term is a tiny fraction, which disappears when taking the floor of $\mathcal{G}_{n,d,k}(t)$. It follows that for virtually all code parameters, the net effect of (21) and (22) consists of eliminating the deficiency term $\Delta(\mathbb{C})$ in Theorems 3 and 4.

III. Discussion and Concluding Remarks

The Han-Siegel probabilistic bound [6, Theorem 3] is the best currently known bound on the stopping redundancy of general linear codes. We presented several improvements upon this bound based upon a more elaborate probabilistic analysis. Although all our results were stated and proved herein for binary linear codes, they extend straightforwardly to linear codes over an arbitrary finite field. Extension to bounds on the stopping redundancy *hierarchy* [8], [9] is also straightforward: simply replace all occurrences of d in Theorems 1–4 and 7 by the corresponding index of the hierarchy. Generalization to bounds on the trapping redundancy [16] should be relatively easy as well, although, perhaps, less straightforward.

While the improvements on the Han–Siegel bound established herein are not dramatic, we believe it is useful to have the strongest possible form of this bound available in the literature. We point out that, in addition to the results presented in Section II, we have investigated several other ideas. For example, one could construct a parity-check matrix for a linear code $\mathbb{C}$ by selecting each codeword of the dual code $\mathbb{C}^{\perp}$ independently with some probability p, and then optimize the value of p. However, we have found that while this method, as well as other probabilistic-choice variants, improve upon the original Han–Siegel bound of [6, Theorem 3], they are generally *less* efficient than the bounds presented in Section II. We note that well-known techniques in "probabilistic method," such as Lovász local lemma [2, p. 64] and Rödl nibble [2, p. 53], do not seem to be applicable in our context. Thus we believe that further improvements, if any, would require drastically new ideas.

Acknowledgment

The authors are grateful to Markus Grassl and Ludo Tolhuizen for their help with references [5] and [10]–[13], respectively.

References

[1] K. A. S. Abdel-Ghaffar and J. H. Weber, "Complete enumeration of stopping sets of full-rank parity-check matrices of Hamming codes," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3196–3201, Sep. 2007.

[2] N. Alon and J. H. Spencer, *The Probabilistic Method*. New York: Wiley, 1991.

[3] C. Di, D. Proietti, I. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1570–1579, Jun. 2002.

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

[5] M. Grassl, *Bounds on the Minimum Distance of Linear Codes,* Jul. 12, 2007 [Online]. Available: at www.codetables.de

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

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

[8] T. Hehn, S. Laendner, O. Milenkovic, and J. B. Huber, "The stopping redundancy hierarchy of cyclic codes," in *Proc. 44th Annu. Allerton Conf. Communications, Control and Computing*, Monticello, IL, Sep. 2006, pp. 1271–1280.

[9] T. Hehn, O. Milenkovic, S. Laendner, and J. B. Huber, "Permutation decoding and the stopping redundancy hierarchy of linear block codes," in *Proc. IEEE Int. Symp. Information Theory*, Nice, France, Jun. 2007, pp. 2926–2930.

[10] H. D. L. Hollmann and L. M. G. M. Tolhuizen, "Generating parity check equations for bounded-distance iterative erasure decoding of even weight codes," in *Proc. 27th Symp. Information Theory in the Benelux*, Noordwijk, The Netherlands, Jun. 2006, pp. 17–24.

[11] H. D. L. Hollmann and L. M. G. M. Tolhuizen, "Generating parity check equations for bounded-distance iterative erasure decoding," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006, pp. 514–517.

[12] H. D. L. Hollmann and L. M. G. M. Tolhuizen, "Generic erasure-correcting sets: bounds and constructions," *J. Combin. Theory, Ser. A*, vol. 113, pp. 1746–1759, Nov. 2006.

[13] 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. Inf. Theory*, vol. 53, no. 2, pp. 823–828, Feb. 2007.

[14] I. N. Kovalenko, "On the limit distribution of the number of solutions of a random system of linear equations in the class of boolean functions," *Theory Probab. App.*, vol. 12, pp. 47–56, 1967.

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

[16] O. Milenkovic, E. Soljanin, and P. Whiting, "Stopping and trapping sets in generalized covering arrays," in *Proc. 40th Conf. Information Science and Systems*, Princeton, NJ, Mar. 2006, pp. 259–264.

[17] E. Rosnes and Ø. Ytrehus, "An algorithm to find all small-size stopping sets of low-density parity-check matrices," in *Proc. IEEE Int. Symp. Information Theory*, Nice, France, Jun. 2007, pp. 975–979.

[18] M. Schwartz and A. Vardy, "On the stopping distance and the stopping redundancy of codes," in *Proc. IEEE Int. Symp. Information Theory*, Adelaide, Australia, Sep. 2005, pp. 975–979.

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

[20] J. H. Weber and K. A. S. Abdel-Ghaffar, "Stopping set analysis for Hamming codes," in *Proc. IEEE Inf. Theory Workshop*, Rotorua, New Zealand, Aug. 2005, pp. 244–247.