

# Group testing schemes from low-weight codewords of BCH codes

Shashanka Ubaru\*

Arya Mazumdar†

Alexander Barg§

**Abstract**—Despite a large volume of research in group testing, explicit small-size group testing schemes are still difficult to construct, and the parameters of known combinatorial schemes are limited by the constraints of the problem. Relaxing the worst-case identification requirements to *probabilistic localization of defectives* enables one to expand the range of parameters, and yet the small-size practical constructions are sparse.

Motivated by this question, we perform an experimental study of almost disjunct matrices constructed from low-weight codewords of binary BCH codes, and evaluate their performance in nonadaptive group testing. We observe that identification of defectives is much more stable in these schemes compared to the schemes constructed from random binary matrices. We derive an estimate of the error probability of identification in the constructed schemes which provides a partial explanation of their performance.

## I. INTRODUCTION

The group testing problem calls for efficient identification of a small number  $t$  of defective elements in population of a large size  $N$ . The elements are tested in groups with the premise that most tests will return negative results, clearing the entire group. If the test result is positive, then the group contains at least one defective. The collection of tests is said to form a *group testing scheme* if the outcomes of the tests enable one to identify any subset of defectives of size at most  $t$ .

A nonadaptive group testing scheme with  $M$  tests is described by an  $M \times N$  binary incidence matrix  $A$ , where each row corresponds to a test, and  $A_{ij} = 1$  if and only if the  $i$ th test includes the  $j$ th element. The result of the test is positive if the indices of ones in the row overlap with the indices of the defective configuration. The smallest possible number of tests in terms of the total number of subjects  $N$  and the maximum number of defective elements  $t$  is known to satisfy  $M = \Theta(\frac{t^2}{\log t} \log N)$  [6]–[8], [20].

A construction of group testing schemes using matrices of error-correcting codes and code concatenation appeared in the foundational paper by Kautz and Singleton [11]. This work introduced a two-level construction in which a  $q$ -ary ( $q > 2$ ) Reed-Solomon code is concatenated with a unit-weight binary code, and the resulting vectors are used as columns of the testing matrix. Many later constructions of group testing schemes

also rely on Reed-Solomon codes and code concatenations; among them [5], [21]. Other explicit constructions of non-adaptive group testing schemes with  $M = O(t^2 \log N)$  were suggested in [10], [18], [19]; see also [3], [4].

In order to improve the tradeoff between the parameters of the scheme, it has been suggested to construct schemes that permit a small probability of error (false positives). Such schemes were considered under the name of *weakly separated designs* in [12], [14], [15], [22]. With this relaxation, it is possible to reduce the number of tests to  $\Theta(t \log N)$  [22]; however, this result is not constructive. A construction of weakly separated designs with  $O(t \text{poly}(\log N))$  was suggested in a recent work [9]. An explicit (non-probabilistic) construction of almost disjunct matrices with the number of tests proportional to  $t^{3/2} \sqrt{\log N}$  was presented in [17] and was subsequently improved to  $t \log^2 N / \log t$  in [16]. The construction of [11] and many others above are based on constant weight error-correcting codes. Estimates of the parameters of the group testing schemes from constant weight codes were obtained using the minimum distance of the code [11] and more recently using the average distance [16], [17]. Another construction of almost disjunct matrices using Reed-Solomon codes was suggested in [2], where it was shown that there exist group testing schemes with  $M = O(t \log N)$  for  $t \leq \log N$ .

In this work, we perform experiments with test matrices constructed from the codewords of a fixed (low) weight in binary BCH codes. The resulting matrices are sparse in the sense that most of their entries are zero. We show that these matrices perform extremely well in experiments, outperforming random matrices. While the experimental results form the main body of the paper, we also give some theoretical justification for this construction. Sharpening the estimates of the error probability is currently an open problem.

A new estimate of the probability of false positive that we derive is based on the dual distance  $d'$  of constant weight codes. Constant weight codes with a given  $d'$  are known as combinatorial designs (of strength  $d' - 1$ ). A design of strength  $r$  (an  $r$ -design, or, in more detail, an  $r$ -( $n, w, \lambda$ ) design) is a collection of  $w$ -subsets of an  $n$ -set  $V$ , called blocks, such that every  $r$  elements of  $V$  are contained in the same number  $\lambda$  of blocks. The use of  $r$ -designs for constructing disjunct matrices is not new, see, e.g., Sect. 7.4 of [3]. However the conclusion in [3, p. 146], is that disjunct matrices obtained from designs are of little interest because of restrictions on their parameters.

The paper is organized as follows. In Section II, we introduce some definitions and notation. The main experimental section appears next in Section III. In Section IV, we present some estimates of the error probability, connecting them with the dual distance of the code. We show that the estimates

\* Dept. of Computer Science, University of Minnesota, Minneapolis, MN 55455. Email: ubaru001@umn.edu. Research supported by NSF grant CCF-1318597.

† Dept. of Computer Science, University of Massachusetts, Amherst, MA 01002. Email: arya@cs.umass.edu. Part of the work was done when the author was at University of Minnesota. Research supported in part by NSF grants CCF1318093 and CCF1453121.

§ Dept. of ECE and ISR, University of Maryland, College Park, MD 20742. Email: abarg@umd.edu. Research supported by NSF grants CCF1217245 and CCF1422955.

available in the literature are quite far from the experimental results, and derive better bounds (which still do not come close to the values observed in our experiments).

## II. DEFINITIONS AND NOTATION

*Definition 1:* An  $M \times N$  binary matrix  $A$  is called  $t$ -disjunct if the support<sup>1</sup> of any of its columns is not contained in the union of the supports of any other  $t$  columns.

It is easy to see that a  $t$ -disjunct matrix gives a group testing scheme that identifies any defective set up to size  $t$ . Conversely, any group testing scheme that identifies any defective set up to size  $t$  must be a  $(t-1)$ -disjunct matrix [3]. To a great advantage, disjunct matrices support a simple identification algorithm that runs in time  $O(Nt)$ . Indeed, any element that participates in a test with a negative outcome is not defective. After we perform all the tests and weed out all the non-defective elements, the disjunctness property of the matrix guarantees that all the remaining elements are defective.

*Definition 2:* For any  $\epsilon > 0$ , an  $M \times N$  matrix  $A$  with columns  $a_1, a_2, \dots, a_N$ , is called  $(t, \epsilon)$ -disjunct if

$$\Pr(\{I \in \binom{[N]}{t}, j \in [N] \setminus I \mid \text{supp}(a_j) \subseteq \bigcup_{k \in I} \text{supp}(a_k)\}) \leq \epsilon.$$

The group testing scheme given by a  $(t, \epsilon)$ -disjunct matrix is called a  $(t, \epsilon)$ -scheme.

In other words, the union of supports of a randomly and uniformly chosen subset of  $t$  columns of a  $(t, \epsilon)$ -disjunct matrix does not contain the support of any other random column with probability at least  $1 - \epsilon$ .

The next fact follows from the definition of disjunct matrices and the decoding procedure [3, p. 134].

*Proposition 1:* A  $(t, \epsilon)$ -disjunct matrix defines a group testing scheme that can identify all items in a random defective configuration of size  $t$ , and with probability  $\epsilon$  identifies any randomly chosen item outside of the defective configuration as defective (false-positive).

*Remark 1:* Unless  $\epsilon < \frac{t}{N-t}$ , the average number of false positives in the  $(t, \epsilon)$  scheme will be greater than the actual number of defectives. However even in that case, the tests will output a subset of  $[N]$  of a vanishingly small proportion that includes all of the  $t$  defective items.

A code of length  $M$  is a subset of the vector space  $\mathbb{F}_q^M$ . The minimum Hamming distance between distinct codewords of  $\mathcal{C}$  is called the *distance of the code*. We use the notation  $\mathcal{C}(M, N, d)$  to refer to the code of length  $M$ , cardinality  $N$  and distance  $d$ . If in addition all the codewords of the code  $\mathcal{C}$  contain exactly  $w$  nonzero entries, we call it a *constant weight code* and use the notation  $\mathcal{C}(M, N, d, w)$ . Finally we refer to a linear code of length  $M$  and dimension  $K$  as an  $[M, K]$  code.

## III. NUMERICAL EXPERIMENTS

As we already mentioned, our work is motivated in part by the experimental results which show that matrices formed of codewords of a fixed weight obtained from binary codes perform very well in the group testing problem for identifying defective entries. In this section, we present a few simulation

results using matrices formed by codewords of fixed weight obtained from different BCH codes, and compare their performance with best possible randomly generated sparse matrices.

The results are presented in Fig. 1,2. Each of the eight plots in Fig. 1 presents results of identification of defectives for two group testing schemes, one using a matrix formed of the fixed-weight codewords of the BCH code and the other using a random binary matrix. The experiments were organized as follows. For instance, for the first plot (top left corner) we formed a test matrix  $A$  by using the codewords of weight  $w = 6$  in the  $[M = 31, K = 21]$  BCH code as its columns. There are  $N = 806$  such codewords in the code, which enables us to construct a  $31 \times 806$  test matrix. This means that we can test a set of  $N = 806$  items for the presence of defectives using  $M = 31$  tests. To compare this construction with the scheme based on random matrices, we constructed a sparse random matrix assigning the entries independently to 0 or 1 with probability of 1 equal to  $\tilde{p} = 1/(t+1)$  and 0 with probability  $1 - \tilde{p}$ . It is known (and also experimentally verified) that  $\tilde{p} = 1/(t+1)$  gives the best performance among such random matrices [3]. Each experiment consisted of generating a random vector with  $t$  defectives randomly inserted among  $N$  items and performing the identification procedure mentioned in Section II above. As the outcome, we record the number of false positives identified by each of the two schemes. This experiment is repeated 300 times; then we compute the average number of false positives found by the two group testing schemes. The curves in the plot show this number as a function of the actual number of defective elements  $t$ .

Similar experiments were performed for the other group testing schemes shown in Fig. 1. In the second plot, we used a  $[63, 57]$ -BCH codeword matrix with constant weight  $w = 3$ ,  $N = 651$  and  $M = 63$ , i.e., the matrix formed by the codewords of weight 3 of the Hamming code of length 63 (they are known to support a 2-design). In the third plot the matrix was formed of codewords of weight 5 in the  $[63, 51]$  BCH code (there are  $N = 1890$  such words) and in the fourth we used the 3411 codewords of weight 7 in the  $[63, 45]$  BCH code. To summarize, the parameters  $(N, M, w)$  of the schemes and the associated codes are:

$$\begin{array}{cccc} (806,31,6) & (651,63,3) & (1890,63,5) & (3411,63,7) \\ [63, 39] & [63,57] & [63,51] & [63,45] \end{array}$$

For the second-row plots in Fig. 1 the parameters are:

$$\begin{array}{cccc} (2170,63,9) & (2667,127,3) & (16002,127,5) & (48387,127,7) \\ [63, 39] & [127,120] & [127,113] & [127,106] \end{array}$$

where we also list the parameters of the BCH codes used to construct the matrices.

We can see that the fixed-weight BCH codeword matrices consistently in most cases perform much better than sparse random matrices in terms of the number of false positives detected, and the gap widens (in most cases) with the increase of the number of defectives. To exemplify this improvement, we show in Fig. 2 two individual experiments performed for fixed-weight BCH testing matrices used to generate Fig. 1, namely, the matrices with the parameters  $N = 1890, M = 63, w = 5$  and  $N = 16002, M = 127, w = 5$ . In the left column in Fig. 2 we show the locations of the actual defective elements inserted in the population and the defective vectors identified

<sup>1</sup>The support of a vector  $x \in \mathbb{F}_q^n$  is the set  $\text{supp}(x) := \{i : x_i \neq 0\}$ .

by the BCH-based matrix (middle plot) and the random matrix (bottom plot) for the first set of parameters. We see that the BCH matrix locates the defectives exactly while the random matrix inserts a large number of false positives compared to the actual number  $t = 4$ . In the right column of the plot we show similar results for an individual experiment for the second set of parameters. Here the BCH matrix-based scheme inserts a few false defectives, while the random matrix adds many more.

#### IV. PROPERTIES OF CONSTANT WEIGHT CODES AND ESTIMATES OF THE ERROR PROBABILITY

In this section we cite some known, and present some new results on constant weight codes with which we attempt to explain the observed performance of constant weight almost disjunct matrices.

The following well-known result of [11] has been the basis of a large number of construction of testing matrices.

*Proposition 2:* An  $(M, N, d, w)$  constant weight binary code  $\mathcal{C}$  provides a  $t$ -disjunct matrix, where  $t = \lfloor \frac{w-1}{w-d/2} \rfloor$ .

*Proof:* Write the codewords of  $\mathcal{C}$  as the columns of an  $M \times N$  matrix. The intersection of supports of any two columns has size at most  $w - d/2$ . Hence if  $w > t(w - d/2)$ , the support of any column will not be contained in the union of supports of any  $t$  other columns. ■

This proposition implies that a group testing scheme can be obtained from constant weight codes with large distance. However, the set of codewords of weight 5 in a  $[63, 51]$  BCH code forms a  $(63, 1890, 3, 5)$  constant weight code with distance  $d = 5$ , so Proposition 2 clearly fails to explain the performance of the group testing scheme obtained from this code.

Extending the theory of Kautz-Singleton to almost-disjunct matrices, [16] provides a bound on the false-positive probability of a constant weight code matrix in terms of its distance distribution. Define the *average distance*  $D$  of a code  $\mathcal{C}$ :

$$D(\mathcal{C}) = \frac{1}{|\mathcal{C}|} \min_{x \in \mathcal{C}} \sum_{y \in \mathcal{C}} d_H(x, y).$$

Here  $d_H$  denotes the Hamming distance. Define also the second-moment of the distance distribution as follows:

$$D_2(\mathcal{C}) = \frac{1}{|\mathcal{C}|^2} \sum_{x, y \in \mathcal{C}} d_H(x, y)^2.$$

One of the main results of [16] is the following theorem.

*Theorem 3:* Let  $\mathcal{C}$  be a constant weight binary code  $\mathcal{C}$  of size  $N$ , minimum distance  $d$  and average distance  $D$  such that every codeword has length  $M$  and weight  $w$ . The test matrix obtained from the code is  $(t, \epsilon)$ -disjunct for the largest  $t$  such that the inequality

$$d \geq D - \frac{3(w - t(w - D/2))^2}{(\ln 1/\epsilon)(2t(w - D/2) + w)}$$

holds true.

Paper [16] also provides a more refined estimate that relies on the second moment of the distance distribution.

*Theorem 4:* Let  $\mathcal{C}$  be a constant-weight  $(M, N, d, w)$  binary code with average distance  $D$  and the second moment of the

TABLE I  
MINIMUM DISTANCE AND AVERAGE DISTANCE

PARAMETERS	$d = d_{\min}$	$D = d_{avg}$	$D_2$
$M = 31, w = 6$	6	9.6894	96.7742
$M = 63, w = 3$	4	5.7231	33.1797
$M = 63, w = 5$	6	9.2112	86.1239
$M = 63, w = 7$	8	12.4481	157.3620
$M = 63, w = 9$	10	15.4357	241.8802
$M = 127, w = 3$	4	5.8605	34.5917
$M = 127, w = 5$	6	9.6050	93.0134
$M = 127, w = 7$	-	-	-

distance distribution  $D_2$ . The test matrix obtained from the code is  $(t, \epsilon)$ -disjunct for the largest  $t$  such that the inequality

$$d \geq D + \frac{3t(D_2 - D^2)}{2(w - t(w - D/2))} - \frac{3(w - t(w - D/2))}{\ln 1/\epsilon} \quad (1)$$

holds true.

In order to compute the estimates based on these theorems, we computed the distance distributions of a number of constant weight codes obtained from low-weight codewords of the BCH codes of length 63 and 127 mentioned above, and found  $D$  and  $D_2$  for these codes. The results are listed in Table I.

Using these values, we can find the estimates of the error probability  $\epsilon$  given by Theorems 3 and 4. The results are summarized in Tables II and III, respectively. Although they represent a large improvement over the initial estimates of Kautz-Singleton, they still do not match the performance in actual experiments. For example, with codewords of weight 5 in a BCH code of length 127, even with 3 defectives the predicted false positive probability is 0.0479, whereas the number of false positives in the experiments is close to zero.

Next we will show that better estimates can be achieved in many cases. To formulate the result we need to define the *dual distance* of a constant weight code. The distance distribution of a constant weight code  $\mathcal{C}(M, N, d, w)$  is a set of numbers  $b_0, b_1, \dots, b_w$ , where

$$b_i = \frac{1}{|\mathcal{C}|} |\{(x, y) \in \mathcal{C}^2 : w - |\text{supp}(x) \cap \text{supp}(y)| = i\}| \quad (2)$$

for  $i = 0, 1, \dots, w$ . Note that  $b_0 = 1$ . The *dual distance*  $d'$  of  $\mathcal{C}$  is defined as

$$d'(\mathcal{C}) = \min \left\{ j \geq 1 : b'_j := \frac{1}{|\mathcal{C}|} \sum_{i=0}^w b_i Q_j(i) > 0 \right\},$$

where  $Q_j(i)$  is the value of the Hahn polynomial of degree  $j$ ; see [13, p. 545].

Now we are ready to state the main theorem in our analysis.

*Theorem 5:* Let  $\mathcal{C}$  be an  $(M, N, d, w)$  constant weight code with dual distance  $d'$  and let  $w < M/2$ . Let  $t$  be the maximum number of defective items and suppose that  $t < M/w$ . For any even  $\ell < d'$  the probability of a false positive test result for the group testing scheme constructed from  $\mathcal{C}$  is bounded above as

$$\epsilon \leq B(\ell, t) \left( \frac{e\ell(M-w)}{2(M-tw)^2} \right)^{\ell/2} \sum_{i=0}^{\ell/2} \left( \frac{(M-w)\ell}{2ew^2} \right)^i, \quad (3)$$

where  $B(\ell, t) = \min\{(18\ell t)^{\ell/2}, t^\ell\}$ . In addition, if  $M \geq \max\{4w^2t/\ell^2, w + 2ew^2/\ell\}$ , then

$$\epsilon \leq t \left( \frac{2\ell^2(M-w)}{(\log \ell)w(M-tw)} \right)^\ell. \quad (4)$$

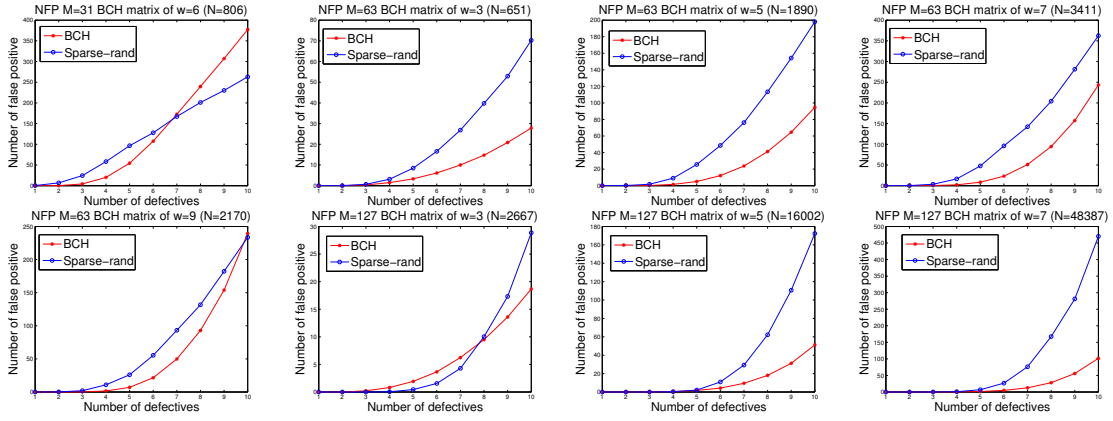


Fig. 1. Number of false positives using various fixed-weight BCH codeword matrices and sparse random matrices with  $\tilde{p} = 1/(t+1)$ , averaged over 300 trials.

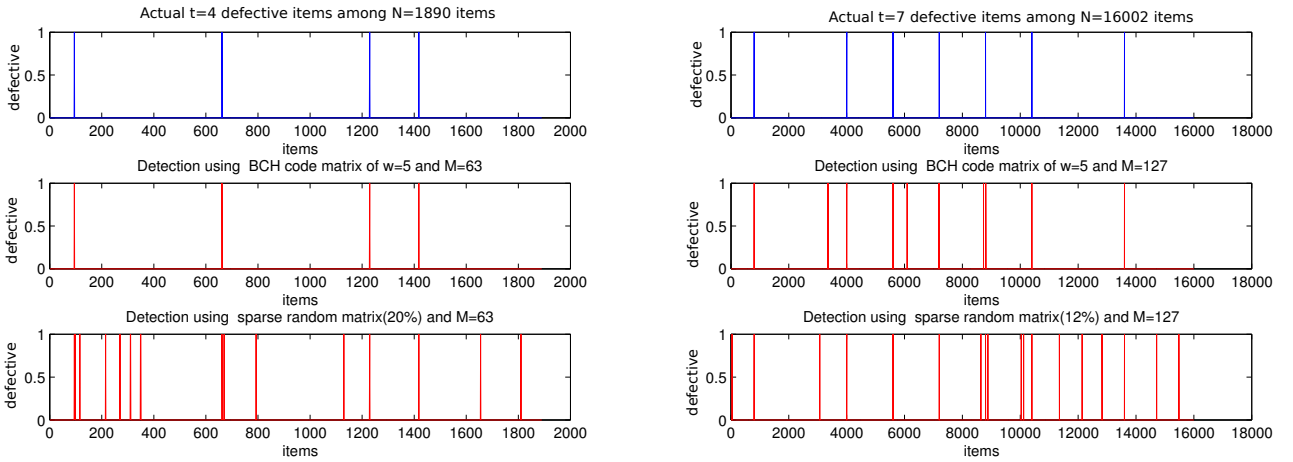


Fig. 2. Two group testing examples. Left: using a fixed-weight BCH codeword matrix with weight  $w = 5$  and a sparse random matrix with  $P(1) = \frac{1}{(t+1)} = 0.20$  for  $N = 1890$  subjects with  $M = 63$  tests. Right: using a fixed-weight  $w = 5$  BCH codeword matrix with  $(N, M) = (16002, 127)$  and a sparse random matrix with  $P(1) = 0.125$ .

TABLE II  
ESTIMATES OF THE ERROR PROBABILITY  $\epsilon$  FROM THEOREM 3

PARAMETERS	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$	$t = 9$	$t = 10$
$M = 31, w = 6$	0.1007	0.3543	0.6707	0.9060	0.9981	0.9630	0.8481	0.6997	0.5500	0.4166
$M = 63, w = 3$	0.0129	0.0267	0.0486	0.0804	0.1228	0.1757	0.2381	0.3084	0.3842	0.4631
$M = 63, w = 5$	0.0326	0.0807	0.1579	0.2619	0.3846	0.5149	0.6419	0.7562	0.8512	0.9232
$M = 63, w = 7$	0.0472	0.1386	0.2847	0.4638	0.6449	0.8007	0.9145	0.9802	1.0000	0.9810
$M = 63, w = 9$	0.0582	0.1979	0.4147	0.6495	0.8428	0.9613	1.0000	0.9716	0.8965	0.7944
$M = 127, w = 3$	0.0122	0.0179	0.0254	0.0350	0.0468	0.0611	0.0778	0.0972	0.1192	0.1438
$M = 127, w = 5$	0.0285	0.0475	0.0733	0.1063	0.1465	0.1933	0.2460	0.3034	0.3642	0.4271

In the case of  $\ell = 2$  we have

$$\epsilon < \frac{t}{M-1} \frac{(M-w)^2}{(M-wt)^2}. \quad (5)$$

*Proof:* (outline) The ideas of the proof are as follows. First we show that as long as  $r < d'$ , the  $r$ th central moment of the distance distribution of the code equals the  $r$ th moment of the hypergeometric random variable with the pmf  $f_X(i) = \binom{w}{i} \binom{M-w}{w-i} / \binom{M}{w}$ ,  $i = 0, 1, \dots, w$ . This fact relies on simple properties of the Johnson association scheme. After that we use a condition similar to  $w > t(w - d/2)$  (see the proof of

Prop. 2) as sufficient for identification and write a Chernov-type bound that it is violated, where the random variables are the indices of  $t$  random columns of the matrix  $A$ . This gives the estimate of  $\epsilon$  in the form of the  $r$ th moment of a sum of certain independent random variables which can be bounded above using classical inequalities such the Rosenthal inequalities or other similar estimates. The full development of these arguments is rather long; see [1].

We have calculated the dual distances of all the low-weight BCH matrices considered in our experiments. The results are summarized in Table V.

TABLE III  
ESTIMATES OF THE ERROR PROBABILITY  $\epsilon$  FROM THEOREM 4

PARAMETERS	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$	$t = 9$	$t = 10$
$M = 31, w = 6$	0.0431	0.1647	0.4311	0.7816	0.9936	0.8957	0.5789	0.2710	0.0928	0.0235
$M = 63, w = 3$	0.0129	0.0264	0.0480	0.0791	0.1206	0.1724	0.2335	0.3024	0.3770	0.4548
$M = 63, w = 5$	0.0226	0.0482	0.0917	0.1578	0.2489	0.3628	0.4929	0.6288	0.7578	0.8673
$M = 63, w = 7$	0.0249	0.0608	0.1302	0.2458	0.4103	0.6078	0.8020	0.9458	1.0000	0.9510
$M = 63, w = 9$	0.0245	0.0728	0.1815	0.3758	0.6409	0.8911	1.0000	0.8955	0.6325	0.3478
$M = 127, w = 3$	0.0122	0.0179	0.0254	0.0349	0.0466	0.0607	0.0774	0.0966	0.1184	0.1428
$M = 127, w = 5$	0.0233	0.0339	0.0479	0.0660	0.0888	0.1168	0.1503	0.1896	0.2346	0.2850

TABLE IV  
ESTIMATES OF THE ERROR PROBABILITY  $\epsilon$  FROM (5), THM. 5

PARAMETERS	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$	$t = 9$	$t = 10$
$M = 63, w = 3$	0.016	0.036	0.060	0.089	0.126	0.172	0.230	0.305	0.403	0.533
$M = 63, w = 5$	0.016	0.039	0.070	0.117	0.188	0.299	0.484	–	–	–
$M = 63, w = 7$	0.016	0.042	0.086	0.165	0.322	0.688	–	–	–	–
$M = 127, w = 3$	0.008	0.017	0.026	0.037	0.049	0.062	0.076	0.092	0.110	0.130
$M = 127, w = 5$	0.008	0.017	0.028	0.041	0.057	0.075	0.098	0.125	0.158	0.199
$M = 127, w = 7$	0.008	0.018	0.030	0.046	0.067	0.095	0.131	0.181	0.251	0.352

TABLE V  
THE DUAL DISTANCES OF BCH CONSTANT WEIGHT CODES

PARAMETERS	$d'$
$n = 31, w = 6$	4
$n = 63, w = 3$	3
$n = 63, w = 5$	3
$n = 63, w = 7$	4
$n = 63, w = 9$	3
$n = 127, w = 3$	3
$n = 127, w = 5$	3

Using this information together with Theorem 5, we can compute upper bounds for the false-positive probabilities obtained from Eq. (5) for the various fixed-weight BCH codeword matrix examples given in Fig. 1. The results are listed in Table IV<sup>2</sup>. We can see that the bounds are small, especially for smaller  $t$ , even when we have  $\ell = 2$ . Compared to all previous results, this gives better estimates of  $\epsilon$ .

## V. CONCLUSION

We propose constructions of almost disjoint matrices from codewords of a fixed weight in binary codes and perform experiments with matrices obtained from BCH codes of length  $M = 63, 127$ . The experiments show that the rate of error (false-positives) of group testing schemes obtained using these matrices is better than the performance of random binary matrices, making them good candidates for practical uses.

We also derive a new estimate of the error probability of constant-weight almost disjoint matrices based on the strength of designs formed by the columns of the matrices. We show that the estimates obtained from our results for the parameters considered are better than the estimates available in the literature. At the same time they still come short of matching the performance of the experiments. Deriving better estimates of the error probability of identification remains an open problem.

## REFERENCES

[1] A. Barg and A. Mazumdar. Almost disjoint matrices from codes and designs. Online at arXiv:1510.02873.

<sup>2</sup> ‘–’ in the table means that the computed estimate is trivial.

- [2] L. Bassalygo and V. Rykov. Multiple-access hyperchannel. *Problems of Information Transmission*, 49(4):399–307, 2013.
- [3] D. Z. Du and F. Hwang. *Combinatorial group testing and its applications*. World Scientific, 2nd edition, 2000.
- [4] D.-Z. Du and F. Hwang. *Pooling designs and nonadaptive group testing*. World Scientific, 2006.
- [5] A. D’yachkov, A. J. Macula, and V. Rykov. New applications and results of superimposed code theory arising from the potentialities of molecular biology. In *Numbers, information and complexity (Bielefeld, 1998)*, pages 265–282. Kluwer Acad. Publ., Boston, MA, 2000.
- [6] A. D’yachkov, I. Vorob’ev, N. Polyansky, and V. Shchukin. Bounds on the rate of disjunctive codes. *Problems of Information Transmission*, 50:27–56, 2014.
- [7] A. G. D’yachkov and V. V. Rykov. Bounds on the length of disjunctive codes. *Problemy Peredachi Informatsii*, 18(3):7–13, 1982.
- [8] Z. Füredi. On  $r$ -cover-free families. *Journ. Combin. Theory, Ser. A*, 73:172–173, 1996.
- [9] A. C. Gilbert, B. Hemenway, A. Rudra, M. J. Strauss, and M. Wootters. Recovering simple signals. In *Information Theory and Applications Workshop (ITA), 2012*, pages 382–391. IEEE, 2012.
- [10] F. Hwang and V. Sós. Non-adaptive hypergeometric group testing. *Studia Sci. Math. Hungar.*, 22:257–263, 1987.
- [11] W. Kautz and R. Singleton. Nonrandom binary superimposed codes. *IEEE Trans. Inform. Theory*, 10(4):363–377, 1964.
- [12] A. J. Macula, V. V. Rykov, and S. Yekhanin. Trivial two-stage group testing for complexes using almost disjoint matrices. *Discrete Applied Mathematics*, 137(1):97–107, 2004.
- [13] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North-Holland, 1977.
- [14] M. B. Malyutov. Mathematical models and results in the theory of screening experiments. In *Problems of Cybernetics*, number 35, pages 5–69. USSR Academy of Sciences, Moscow, 1977. in Russian.
- [15] M. B. Malyutov. The separating property of random matrices. *Mathematical Notes*, 23(1):84–91, 1978.
- [16] A. Mazumdar. Nonadaptive group testing with random set of defectives via constant-weight codes. arXiv:1503.03597.
- [17] A. Mazumdar. On almost disjoint matrices for group testing. In *Algorithms and Computation*, pages 649–658. Springer, 2012.
- [18] Q. A. Nguen and T. Zeisei. Bounds on constant weight binary superimposed codes. *Probl. Contr. Inform. Theory*, 17:223–230, 1988.
- [19] E. Porat and A. Rothschild. Explicit non-adaptive combinatorial group testing schemes. In *Automata, languages and programming. Part I*, volume 5125 of *Lecture Notes in Comput. Sci.*, pages 748–759. Springer, Berlin, 2008.
- [20] M. Ruszinkó. On the upper bound on the size of the  $r$ -cover-free families. *Journal of Combinatorial Theory, Series A*, 66:302–310, 1994.
- [21] V. M. Sidelnikov and O. Y. Prikhodov. On the construction of  $(w, r)$  cover-free codes. *Probl. Inform. Trans.*, 45(1):36–40, 2009.
- [22] A. Zhigljavsky. Probabilistic existence theorems in group testing. *Journal of statistical planning and inference*, 115(1):1–43, 2003.