ON FINDING A SUBSET OF NON-DEFECTIVE ITEMS FROM A LARGE POPULATION USING GROUP TESTS: RECOVERY ALGORITHMS AND BOUNDS

Abhay Sharma and Chandra R. Murthy

Dept. of ECE, Indian Institute of Science, Bangalore, India

ABSTRACT

We present computationally efficient and analytically tractable algorithms for identifying a given number of "non-defective" items from a large population containing a small number of "defective" items under a noisy Non-adaptive Group Testing (NGT) framework. In contrast to the classical NGT, where the main goal is to identify the *complete* set of defective items, the main goal of a *non-defective subset recovery* algorithm is to identify a *subset* of non-defective items given the test outcomes. In this paper, we present three algorithms and corresponding bounds on the number of tests required for successful non-defective subset recovery. We consider a random, non-adaptive pooling strategy with noisy test outcomes, where we account for the impact of both additive noise (false positives) and dilution noise (false negatives). We provide simulation results to highlight the relative performance of the algorithms, and to demonstrate the significant improvement they offer over existing approaches, in terms of the number of tests required for a given success rate.

Index Terms— Healthy subset identification, Finding zeros, Non-adaptive group testing, Data streaming, Medical screening.

1. INTRODUCTION

Group testing finds applications in diverse engineering fields such as DNA sequencing, medical screening [1], data streaming and sketching [2, 3], industrial testing [4], data pattern mining [5] etc. The basic goal in classical group testing is to identify a small set of K unknown "defective" items from a large set of N items by performing a relatively small number of group tests [1]. Each group test provides a binary indication as to whether or not the pool of items under test contains any defective items. One of the useful variants of group testing is non-adaptive group testing (NGT) [1, 6, 7], where different tests are conducted simultaneously, i.e., the tests do not use information provided by the outcome of any other test. An important aspect of NGT is the design of the groups or pools of individuals that go into each test. One popular approach is random pooling [6, 8, 9], where the items included in the group test are chosen uniformly at random from the population. With random pooling, a key issue is the design of computationally efficient recovery algorithms for the defective set, given the set of noisy test outcomes. In this work, in contrast to the defective set identification problem, we study the *non-defective subset identification* problem in the noisy, non-adaptive group testing with random pooling (NNGT-R) setup.

We refer to the non-defective subset identification problem as that of finding a *subset* consisting of L $(\leq N-K)$ non-defective items from a population of N items containing a set of $K \ll N$ defective items [10]. There are many applications where the goal is to identify only a small subset of non-defective items rather than identifying all the defective items. For example, consider the spectrum hole search problem in a cognitive radio (CR) network setup. It is known that the primary user occupancy is sparse in the frequency domain, over a wide band of interest [11, 12]. This is equivalent to having a small subset of defective items embedded in a large set of candidate frequency bins. The secondary users do not need to identify all the frequency bins occupied by the primary users; they only need to discover relatively small unoccupied sub-bands to setup the secondary communications, i.e., a non-defective subset identification problem. As another example, consider a scenario from the data stream domain [2, 3]. We receive a high volume SMS data stream in response to a trivia contest run during a television show. The SMS data is processed to ascertain whether the answer is correct. The outcome is streamed to the TV studio server as $\langle$ *phone.number, flag* $\rangle$, where *flag*$= 1 (= -1)$ indicating a correct (wrong) answer. Owing to the simplicity of trivia questions, we expect a large majority of the *flag* variables to be equal to 1. Due to large number of received records and severe memory constraints, the data stream is often summarized using a small number of "sketches" using test matrices, and the sketch vector is equivalent to the outcome vector in the NGT setting [2, 9]. The objective is to use the sketch vector to identify a small group of responders with correct answers, i.e., the winners of the contest, and is thus a non-defective subset identification problem. In [10], using information theoretic arguments, it was shown that compared to the conventional approaches of identifying the non-defective subset by first identifying the defective set or by testing individual items one by one, directly searching for an L-sized non-defective subset offers a significant reduction in the number of tests, especially when L is small compared to $N - K$. In this paper, we develop computationally efficient algorithms for non-defective subset identification in an NNGT-R framework.

Although the problem of non-defective subset identification has not yet been explored in the literature, it is a generalization of the defective set identification problem. In particular, notice that identifying $L = N - K$ non-defective items is equivalent to identifying K defective items. Hence, the algorithms presented in this work can be related to algorithms from the rich available literature for finding the defective set; see [1] for an excellent collection of existing results and references. In general, for the NNGT-R framework, three broad approaches have been adopted for defective set recovery [7]. First, the row based approach, also frequently referred to as the "naive" decoding algorithm, finds the defective set by finding *all* the non-defective items [7, 13]. The second popular decoding approach is based on the idea of finding defective items iteratively (or greedily) by "appropriately" matching the column of the test matrix corresponding to a given item with the test outcome vector [1, 6, 7, 14]. A recent work, [15], investigates the problem of finding zeros in a sparse vector in the compressive sensing framework, and also proposes a greedy algorithm based on correlating the columns of sens-

The financial support of the Department of Electronics and Information Technology, Govt. of India, is gratefully acknowledged.

ing matrix with the output vector (i.e., column matching). Finally, linear programming relaxation based algorithms have been proposed in [7, 16] for defective set identification in group testing. A class of linear programs is setup by letting the boolean variables take real values (between 0 and 1) and imposing inequality or equality constraints to model the outcome of each pool.

In this work, we propose novel algorithms for identifying a non-defective subset in an NNGT-R framework. We derive non-asymptotic upper bounds on the average error rate that lead to a theoretical guarantee on the number of tests for the proposed algorithms. We summarize our main contributions as follows:

- We propose three computationally efficient and analytically tractable algorithms for identifying a non-defective subset of given size in a NNGT-R framework (see Section 3): **RoAl** (row based algorithm), **CoAl** (column based algorithm) and **RoLpAl** (LP relaxation based algorithm).
- We derive bounds on the number of tests that guarantee successful non-defective subset recovery for each algorithm. The derived bounds are a function of the system parameters, namely, the number of defective items, the size of non-defective subset, the population size and the noise parameters.

We present numerical simulations to compare the relative performance of the algorithms and to illustrate the advantage of the proposed algorithms compared to the conventional methods based on identifying the defective set followed by picking the required number of items from the complement set (Section 4). Due to lack of space, the proofs have been omitted; these will be presented in a journal version of this work.

Notation: For any positive integer n, $[n] \triangleq \{1, 2, \ldots, n\}$. For a vector $\underline{a}$, $\underline{a}(i)$ denotes its i^{th} component. supp$(\underline{a})$ denotes the support set for the vector $\underline{a}$. In the context of boolean vectors, $\underline{a}^c$ denotes the component wise boolean complement of $\underline{a}$. $\underline{1}_n$ and $\underline{0}_n$ denote an all-one and all-zero vector, respectively, of size n. $\underline{a} \preccurlyeq \underline{b}$ denotes the component-wise inequality, i.e., it means $\underline{a}(i) \leq \underline{b}(i)\ \forall i$. $\mathcal{B}(q), q \in [0\ 1]$ denotes the Bernoulli distribution with parameter q.

2. SIGNAL MODEL

In our setup, we have a population of N items, out of which K are defective. Let $\mathcal{G} \subset [N]$ denote the defective set, such that $|\mathcal{G}| = K$. We consider a non-adaptive group testing framework with random pooling designs [1, 7, 17, 18], where all the group tests are decided a priori and the items to be pooled in a given test are chosen randomly. The group tests are defined by a boolean matrix, $\mathbf{X} \in \{0, 1\}^{M \times N}$, that assigns different items to the M group tests (pools). The j^{th} pool tests the items corresponding to the columns with 1 in the j^{th} row of $\mathbf{X}$. We consider an i.i.d. random Bernoulli test matrix [17], where each $X_{ij} \sim \mathcal{B}(p)$ for some $0 < p < 1$. Thus, M randomly generated pools are specified. In the above, p is a design parameter that controls the average number of items being tested in a single group test. In particular, we choose $p = \frac{\alpha}{K}$, and a specific value of α is chosen based on the analysis of different algorithms.[1]

When the tests are completely reliable, then the output of the M tests is given by the boolean OR of the columns of $\mathbf{X}$ corresponding to the defective set $\mathcal{G}$. In group testing, two different noise models are considered [6, 7, 17]: (a) An *additive* noise model, where there is a probability, $q \in (0, 0.5]$, that the outcome of a group test containing only non-defective items turns out to be positive (Fig. 1); (b) A *dilution* model, where there is a probability, $u \in (0, 0.5]$, that a given item does not participate in a given group test (Fig. 1). Let $\underline{d}_i \in \{0, 1\}^M$. Let $\underline{d}_i(j) \sim \mathcal{B}(1 - u)$ be chosen independently for all $j = 1, 2, \ldots, M$ and for all $i = 1, 2, \ldots, N$. Let $\mathbf{D}_i \triangleq \text{diag}(\underline{d}_i)$. Let "$\bigvee$" denote the boolean OR operation. The output vector $\underline{y} \in \{0, 1\}^M$ can be represented as

$$\underline{y} = \bigvee_{i \in \mathcal{G}} \mathbf{D}_i \underline{x}_i \bigvee \underline{w}, \tag{1}$$

where $\underline{x}_i \in \{0, 1\}^M$ is the i^{th} column of $\mathbf{X}$, $\underline{w} \in \{0, 1\}^M$ is the additive noise with the i^{th} component $\underline{w}(i) \sim \mathcal{B}(q)$. Given the test output vector, $\underline{y}$, our goals are the following:

(a) To find computationally efficient algorithms to identify L non-defective items, i.e., an L-sized subset belonging to $[N] \backslash \mathcal{G}$.

(b) To analyze the performance of the proposed algorithms with the objective of finding the number of tests required for non-defective subset recovery with high probability of success.

As is common in the literature for defective set recovery in group testing or sparse vector recovery in compressed sensing, there exist two types of recovery results: (a) *Non-uniform/Per-Instance recovery results*: These state that a randomly chosen test matrix leads to successful non-defective subset recovery with high probability of success for a given fixed defective set, and, (b) *Uniform/Universal recovery results*: These state that a random draw of test matrix leads to a successful non-defective subset recovery with high probability of success for all possible defective sets. It is possible to easily extend non-uniform results to the uniform case using union bounds. Hence, we primarily focus on the non-uniform recovery results and present the extension to the uniform case for one of the proposed algorithms (see Corollary 1).

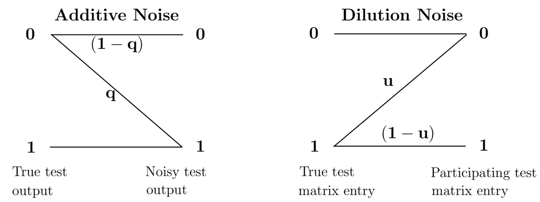

Fig. 1: Different types of noise in the group testing signal model.

3. ALGORITHMS AND MAIN RESULTS

We now propose three algorithms for non-defective subset recovery. Each algorithm takes the observed noisy output vector $\underline{y} \in \{0, 1\}^M$ and the test matrix $\mathbf{X} \in \{0, 1\}^{M \times N}$ as inputs, and outputs a set of L non-defective items, $\hat{S}_L$. The recovery is successful if the declared set does not contain any defective item, i.e., $\hat{S}_L \cap S_d = \{\emptyset\}$.

3.1. Row Based Algorithm

Our first algorithm to find non-defective items is also the simplest and the most intuitive one. We make use of the basic fact of group testing that, in the noiseless case, if the test outcome is negative, then all the items being tested are non-defective.

[1]The above parametrized form of p is motivated by our earlier work [10], where one of the conclusions, based on information theoretic arguments, was that the optimal value of p that minimizes the number of tests required for "finding a non-defective subset" and "finding the defective set" are the same. The form α/K approximates this optimal value very well and has been widely used in literature [6, 7, 19] when probabilistic constructions are employed for designing the test matrices.

RoAl (Row based algorithm):

- Compute $\underline{z} = \sum_{j \in \text{supp}(\underline{y}^c)} \underline{x}_j^{(r)}$, where $\underline{x}_j^{(r)}$ is the j^{th} row of the test matrix.
- Order the entries of $\underline{z}$ in descending order.
- Declare the items indexed by the top L entries as the non-defective subset.

That is, declare the L items that have been tested most number of times in pools with negative outcomes as non-defective items. Recall that dilution noise can lead to a test containing defective items in the pool being declared negative, resulting in a possible mis-classification of the defective items. On the other hand, since the algorithm only considers tests with negative outcomes, additive noise does not lead to mis-classification of defective items as non-defective. However, the additive noise does lead to an increased number of tests as the algorithm might have to discard many of the pools that contain only non-defective items.

Note that existing row based algorithms for finding the defective set [1, 7] can be obtained as a special case of **RoAl** by setting $L = N - K$, i.e., by looking for all non-defective items. However, the analysis in the past work does not quantify the impact of parameter L, and that is our main goal here. We characterize the number of tests, M, that are required to find L non-defective items with high probability of success using **RoAl** in the following theorem:

Theorem 1. *(Non-Uniform recovery with **RoAl**) Let N, L, M, p, u and q be as defined above. Define $\gamma_0 \triangleq \frac{u}{(1-(1-u)p)}$. Let p be chosen as $\frac{\alpha}{K}$ with $\alpha = \frac{1}{3(1-u)}$. There exist absolute constants $C_0 > 0$ such that, if the number of tests are chosen as*

$$M \geq \frac{C_0 K(1-u)}{(1-q)(1-\gamma_0)^2} \left(\frac{\log \left[K \binom{N-K}{L-1} \right]}{(N-K)-(L-1)} \right), \qquad (2)$$

*then for a given defective set there exist positive constants c_0, c_1, such that the algorithm **RoAl** finds L non-defective items with probability exceeding $1 - \exp(-Mc_0) - \exp(-Mc_1)$.*

The following corollary extends the above result to uniform recovery of a non-defective subset using **RoAl**.

Corollary 1. *(Uniform recovery with **RoAl**) Let p and γ_0 be as defined in Theorem 1. Let $N_0 \triangleq (N-K)-(L-1)$. There exist absolute constants $C_0 > 0$ and $C_1 > 0$ such that, if the number of tests are chosen as*

$$M \geq \max \left\{ \frac{C_0 K(1-u)}{(1-q)(1-\gamma_0)^2} \frac{\log \left[N \binom{N-K}{L-1} \right]}{N_0}, \frac{C_1 \log \binom{N}{K}}{(1-q)} \right\},$$

*then for **any** defective set there exist positive constants $c_0, c_1 > 0$ such that the **RoAl** finds L non-defective items with probability exceeding $1 - \exp(-Mc_0) - \exp(-Mc_1)$.*

3.2. Column Based Algorithm

The column based algorithm is based on matching the columns of the test matrix with the outcome vector. A non-defective item does not impact the output and hence the corresponding column in the test matrix should be "uncorrelated" with the output. On the other hand, "most" of the pools that test a defective item should test positive. This forms the basis of distinguishing a defective item from a non-defective one. The specific algorithm is as follows:

CoAl (Column based algorithm): Let $\psi_{cb} > 0$ be some constant.

- Compute $\mathcal{T}(i) = \underline{x}_i^T \underline{y}^c - \psi_{cb}(\underline{x}_i^T \underline{y})$ for each $i = 1, \ldots, N$, where $\underline{x}_i$ is the i^{th} column of $\mathbf{X}$.
- Sort $\mathcal{T}(i)$ in descending order.
- Declare the items indexed by top L entries as the non-defective subset.

We note that, in contrast to the row based algorithm, **CoAl** works with pools of both negative and positive test outcomes. In the above algorithm, the constant ψ_{cb} can be tuned for best performance. In the sequel, we set ψ_{cb} to be the value that optimizes an upper bound on the number of tests, as presented in the following theorem:

Theorem 2. *(Non-Uniform recovery with **CoAl**) Let N, L, M, p, u and q be as defined above. Let $\Gamma \triangleq (1-q)\left(1-(1-u)p\right)^K$ and $\gamma_0 \triangleq \frac{u}{(1-(1-u)p)}$. Let p be chosen as $\frac{1}{3(1-u)K}$. Let $\psi_0 \triangleq \frac{\Gamma(1+\gamma_0)}{2(1-p)}$ and choose $\psi_{cb} = \psi_0$. There exists absolute constant $C_2 > 0$ such that, if the number of tests are chosen as*

$$M \geq \frac{C_2 K(1-u)}{(1-\gamma_0)^2(1+\psi_0)(1-q)} \left(\frac{\log \left[K \binom{N-K}{L-1} \right]}{(N-K)-(L-1)} \right), \quad (3)$$

*then for a given defective set there exists $c_0 > 0$ such that **CoAl** finds L non-defective items with probability exceeding $1 - \exp(-Mc_0)$.*

It is tempting to compare the performance of **RoAl** and **CoAl** by comparing required number of tests in (2) and (3), respectively. However, such comparisons must be done with care, keeping in mind that the required number of observations in (2) and (3) are based on an upper bound on the average probability of error. The main objective of these results is to provide a guarantee on the number of tests required for non-defective subset recovery and highlight the order-wise dependence of the number of tests on the system parameters. For the comparison of the relative performance of the algorithms, we refer the reader to Section 4, where we present numerical results obtained from simulations.

3.3. Linear program relaxation based algorithm

In this section, we consider linear program (LP) relaxations to the non-defective subset recovery problem and identify the conditions under which such LP relaxations lead to recovery of a non-defective subset with high probability of success. Let Y_z be the set of all the pools with negative outcomes and $M_z \triangleq |Y_z|$. Let $\mathbf{X}(Y_z, :)$ denote the $M_z \times N$ sized sub-matrix containing only the rows indexed by Y_z. Define the following linear program, with optimization variables $\underline{z} \in \mathbb{R}^N$ and $\underline{\eta} \in \mathbb{R}^{M_z}$:

$$\underset{\underline{z},\underline{\eta}}{\text{minimize}} \quad \underline{1}_{M_z}^T \underline{\eta} \qquad (4)$$

$$\textbf{(LP0)} \qquad \text{subject to} \quad \mathbf{X}(Y_z,:)(\underline{1}_N - \underline{z}) - \underline{\eta} = \underline{0}_{M_z}, \qquad (5)$$

$$\underline{0}_N \preccurlyeq \underline{z} \preccurlyeq \underline{1}_N, \;\; \underline{\eta} \succcurlyeq \underline{0}_{M_z},$$

$$\underline{1}_N^T \underline{z} \leq L.$$

RoLpAl (LP relaxation with negative outcome pools only)

- Setup and solve **LP0**. Let $\underline{\hat{z}}$ be the solution of **LP0**.
- Sort $\underline{\hat{z}}$ in descending order.
- Declare the items indexed by the top L entries as the non-defective subset.

The above program relaxes the combinatorial problem of choosing L out of N items by allowing the boolean variables to acquire "real" values between 0 and 1 as long as the constraints imposed by negative pools, specified in (5), are met. Intuitively, the variable $\underline{z}$ (or the variable $[\underline{1}_N - \underline{z}]$) can be thought of as the confidence with which an item is declared as non-defective (or defective). The constraint $\underline{1}_N^T \underline{z} \leq L$ forces the program to assign high values (close to 1) for "approximately" the top L entries only, which are then declared as non-defective. The error analysis proceeds by first deriving sufficient conditions for the non-defective subset recovery with **RoLpAl** in terms of the dual variables of **LP0**. We then derive the number of tests required to satisfy these sufficiency conditions with high probability. We summarize the main result in the following theorem:

Theorem 3. *(Non-Uniform recovery with* ***RoLpAl*** *) Let N, L, M, p, u and q be as defined above. Let p be chosen as $\frac{1}{3(1-u)K}$. If the number of tests are chosen as (2), then for a given defective set, there exist positive constants c_0, c_1, such that* ***RoLpAl*** *finds L non-defective items with probability exceeding* $1 - \exp(-Mc_0) - \exp(-Mc_1)$.

4. SIMULATIONS

In this section, we investigate the empirical performance of the algorithms for non-defective subset recovery proposed in this work. In contrast to the previous section, where theoretical guarantees on the number of tests were derived based on the analysis of the upper bounds on probability of error of these algorithms, here we empirically find the exact number of tests required to achieve a given performance level. Our setup is as follows. For a given set of operating parameters, i.e., N, K, u, q and M, we choose a defective set $S_d \subset [N]$ randomly such that $|S_d| = K$, and generate the test output vector $\underline{y}$ according to (1). We then recover a subset of L non-defective items using different recovery algorithms, i.e., **RoAl**, **CoAl** and **RoLpAl**, and compare it with the defective set. This experiment is repeated for different values of M and L. For each trial, the test matrix $\mathbf{X}$ is generated with random Bernoulli i.i.d. entries, i.e., $\mathbf{X}_{ij} \sim \mathcal{B}(p)$, where p is a design parameter. We choose $p = \frac{1}{K}$ for the reasons mentioned earlier. Also, for **CoAl**, as suggested by Theorem 2, we set $\psi_{cb} = \frac{\Gamma(1+\gamma_0)}{2(1-p)}$. Unless otherwise stated, we set $N = 256$, $K = 16$, $u = 0.05$, $q = 0.1$ and vary L and M.

Figure 2 shows the variation of the empirical probability of error with the number of tests, for $L = 64$ and $L = 128$. These curves demonstrate the theoretically expected exponential behavior of the average error rates and the similarity of the error rate performance of **RoAl** and **RoLpAl**. We also note that, as expected, the algorithm that use tests with both positive and negative outcomes perform better than the algorithms that use only tests with negative outcomes. Figure 3 presents the number of tests M required to achieve a target error rate of 10% as a function of the *size* of non-defective subset, L. We note that, for small values of L, all algorithms perform similarly, but, in general, **CoAl** is the best performing algorithm across all values of L. We also compare the algorithms proposed in this work with the indirect approach of identifying the non-defective items by first

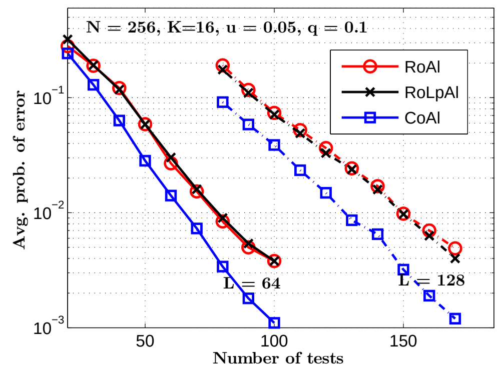

Fig. 2: Average probability of error (APER) *vs.* number of tests M for all algorithms. The APER decays exponentially with M.

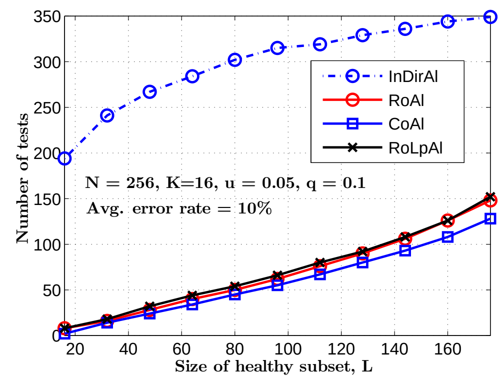

Fig. 3: Number of tests *vs.* size of the non-defective subset. Algorithm **CoAl** performs the best among the ones considered. The direct approach for finding non-defective items outperforms the indirect approach ("InDirAl") [10].

identifying the defective items [10]. We first employ a defective set recovery algorithm for identifying a defective set and then choose L items uniformly at random from the complement set. This algorithm is referred to as "InDirAl" algorithm in Figure 3. In particular, we have used the "**No-LiPo-**" [7] for defective set identification. It can be easily seen that the "direct" approach significantly outperforms the "indirect" approach.

5. CONCLUSIONS

In this work, we proposed analytically tractable and computationally efficient algorithms for identifying a non-defective subset of a given size in a noisy non-adaptive group testing setup. We presented upper bounds on the number of tests for guaranteed correct identification. Also, it was found that the column based algorithm **CoAl** gave the best performance for a wide range of values of L, the size of non-defective subset to be identified. In this work, we have considered a randomized pooling strategy. An interesting problem for future work is to devise deterministic group test constructions for the purpose of non-defective subset identification.

REFERENCES

[1] D. Du and F. Hwang, *Pooling designs and non-adaptive group testing: Important tools for DNA sequencing*. World Scientific, 2006.

[2] A. C. Gilbert and M. J. Strauss, “Analysis of data streams: Computational and algorithmic challenges,” *Technometrics*, vol. 49, no. 3, pp. 346–356, 2007.

[3] G. Cormode and S. Muthukrishnan, “What’s hot and what’s not: Tracking most frequent items dynamically,” *ACM Trans. Database Syst.*, vol. 30, no. 1, pp. 249–278, Mar. 2005. URL: http://doi.acm.org/10.1145/1061318.1061325

[4] M. Sobel and P. A. Groll, “Group testing to eliminate efficiently all defectives in a binomial sample,” *Bell System Technical Journal*, vol. 38, no. 5, pp. 1179–1252, 1959.

[5] A. J. Macula and L. J. Popyack, “A group testing method for finding patterns in data,” *Discrete Appl. Math.*, vol. 144, no. 1-2, pp. 149–157, 2004. URL: http://dx.doi.org/10.1016/j.dam.2003.07.009

[6] M. Cheraghchi, A. Hormati, A. Karbasi, and M. Vetterli, “Group testing with probabilistic tests: Theory, design and application,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 7057–7067, Oct. 2011.

[7] C. L. Chan, S. Jaggi, V. Saligrama, and S. Agnihotri, “Non-adaptive group testing: Explicit bounds and novel algorithms,” *CoRR*, vol. abs/1202.0206, 2012.

[8] A. Sebo, “On two random search problems,” *Journal of Statistical Planning and Inference*, vol. 11, no. 1, pp. 23–31, Jan. 1985.

[9] A. C. Gilbert, M. A. Iwen, and M. J. Strauss, “Group testing and sparse signal recovery,” in *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, Oct. 2008, pp. 1059–1063.

[10] A. Sharma and C. R. Murthy, “On finding a subset of healthy individuals from a large population,” *CoRR*, vol. abs/1307.8240, 2013.

[11] D. Cabric, S. M. Mishra, D. Willkomm, R. Brodersen, and A. Wolisz, “A cognitive radio approach for usage of virtual unlicensed spectrum,” in *Proc. of 14th IST Mobile Wireless Communications Summit*, 2005.

[12] A. Sharma and C. R. Murthy, “Group testing based spectrum hole search for cognitive radios,” *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3794–3805, Oct. 2014.

[13] H.-B. Chen and F. K. Hwang, “A survey on nonadaptive group testing algorithms through the angle of decoding,” *Journal of Combinatorial Optimization*, vol. 15, no. 1, pp. 49–59, 2008.

[14] M. Aldridge, L. Baldassini, and O. Johnson, “Group testing algorithms: Bounds and simulations,” *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3671–3687, June 2014.

[15] J. Yoo, Y. Xie, A. Harms, W. U. Bajwa, and R. A. Calderbank, “Finding zeros: Greedy detection of holes,” *CoRR*, vol. abs/1303.2048, 2013.

[16] D. M. Malioutov and M. B. Malyutov, “Boolean compressed sensing: LP relaxation for group testing.” in *Proc. ICASSP*, 2012, pp. 3305–3308.

[17] G. Atia and V. Saligrama, “Boolean compressed sensing and noisy group testing,” *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1880–1901, 2012.

[18] M. B. Malyutov, “The separating property of random matrices,” *Mat. Zametki*, vol. 23, no. 1, pp. 155–167, 1978. URL: http://dx.doi.org/10.1007/BF01104893

[19] S. Aeron, V. Saligrama, and M. Zhao, “Information theoretic bounds for compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5111 –5130, Oct. 2010.