Sparse Graph Codes for Non-adaptive Quantitative Group Testing

Esmaeil Karimi, Fatemeh Kazemi, Anoosheh Heidarzadeh, Krishna R. Narayanan, and Alex Sprintson

***Abstract*— This paper considers the problem of Quantitative Group Testing (QGT). Consider a set of N items among which K items are defective. The QGT problem is to identify (all or a sufficiently large fraction of) the defective items, where the result of a test reveals the number of defective items in the tested group. In this work, we propose a non-adaptive QGT scheme using sparse graph codes over bi-regular bipartite graphs and binary t-error-correcting BCH codes. The proposed scheme provides exact recovery with probabilistic guarantee, i.e. recovers all the defective items with high probability. In particular, we show that for the sub-linear regime where $\frac{K}{N}$ vanishes as $K, N \to \infty$, the proposed scheme requires at most $m \approx 1.19K \log_2 \left(4.74 \frac{N}{K}\right)$ tests to recover all the defective items with probability approaching one as $K, N \to \infty$. This bound can be achieved by $t = 2$. The testing and recovery algorithms of the proposed scheme for any $t \leq 4$ have the computational complexity of $\mathcal{O}(K \log^2 \frac{N}{K})$ and $\mathcal{O}(K \log \frac{N}{K})$, respectively. Our simulation results also show that the proposed scheme significantly outperforms a non-adaptive semi-quantitative group testing scheme recently proposed by Abdalla *et al.* in terms of the required number of tests for identifying all the defective items with high probability.**

I. INTRODUCTION

In this work, we consider the problem of Quantitative Group Testing (QGT). Consider a set of N items among which K items are defective. The QGT problem is to identify (all or a sufficiently large fraction of) the defective items, where the result of a test reveals the number of defective items in the tested group. The key difference between the QGT problem and the original group testing problem is that, unlike the former, in the latter the result of each test is either 1 or 0 depending on whether the tested group contains any defective items or not. The objective of QGT is to design a test plan with minimum number of tests that identifies (all or a sufficiently large fraction of) the defective items.

There are two general categories of test strategies: *non-adaptive* and *adaptive*. In an adaptive scheme, each test depends on the outcomes of the previous tests. On the other hand, in a non-adaptive scheme, all tests are planned in advance. In other words, the result of one test does not affect the design of another test. Not only are order-optimal non-adaptive algorithms as effective as order-optimal adaptive algorithms asymptotically, but also in most practical applications non-adaptive algorithms are preferred since they allow one to perform all tests at once in parallel.

The authors are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (E-mail: {esmaeil.karimi, fatemeh.kazemi, anoosheh, krn, spalex}@tamu.edu).

This material is based upon work supported by the National Science Foundation under Grants No. 1718658, 1642983, and 1547447.

Let S be the index set of the defective items and $\hat{S}$ be an estimation of S. Depending on the application at hand, there can be different requirements for the *closeness* of $\hat{S}$ to S [1], [2]. The strongest condition for closeness is *exact recovery* when it is required that $\hat{S} = S$. Two weaker conditions are *partial recovery without false detections* when it is required that $\hat{S} \subseteq S$ and $|\hat{S}| \geq (1-\epsilon)|S|$, and *partial recovery without missed detections* when it is required that $S \subseteq \hat{S}$ and $|\hat{S}| \leq (1+\epsilon)|S|$. There are also different types of the *recovery guarantees* [2]. The strongest guarantee is *perfect recovery guarantee* when the exact or partial recovery needs to be achieved with probability 1 (over the space of all problem instances). A slightly weaker guarantee is *probabilistic recovery guarantee* when it suffices to achieve the exact or partial recovery with high probability only (and not necessarily with probability 1). In this work, we are interested in the exact recovery of all defective items with the probabilistic recovery guarantee.

A. Related Work and Applications

The QGT problem has been extensively studied for a wide range of applications, e.g., multi-access communication, spectrum sensing, and network tomography, see, e.g., [3], and references therein. This problem was first introduced by Shapiro in [4]. Several non-adaptive and adaptive QGT strategies have been previously proposed, see, e.g., [3], [5], [6]. It was shown in [7] that any non-adaptive algorithm must perform at least $(2K \log_2(N/K))/\log_2 K$ tests. Various order optimal or near-optimal non-adaptive strategies were previously proposed, see, e.g., [5]–[8]. In particular, the work of [8] gives a non-constructive probabilistic proof for the existence of an order-optimal non-adaptive algorithm. The best known polynomial-time non-adaptive algorithms require $K \log_2 N$ tests [7]. Recently, a semi-quantitative group testing scheme based on sparse graph codes was proposed in [9], where the result of each test is an integer in the set $\{0, 1, 2, \ldots, L\}$. This strategy identifies a $(1-\epsilon)$ fraction of defective items using $c(\epsilon, L)K \log_2 N$ tests with high probability, where $c(\epsilon, L)$ depends only on ϵ and L.

B. Connection with Compressed Sensing

A closely related problem to QGT is the problem of compressed sensing (CS) in which the goal is to recover a sparse signal from a set of (linear) measurements. Given an N-dimensional sparse signal with a support size up to K, the objective is to identify the indices and the values of non-zero elements of the signal with minimum number of

978-1-5386-6900-6/19/$31.00 ©2019 IEEE

measurements. The main differences between the CS problem and the QGT problem are in the signal model and the constraints on the measurement matrix. Most of the existing works on the CS problem consider real-valued signals and measurement matrices. The QGT problem, however, requires both the signal and the measurement matrix to be binary.

There are a number of CS algorithms in the literature that use binary measurement matrices, see, e.g. [10], [11] and references therein. However, these strategies either use techniques not applicable to binary signals, or provide different types of closeness and guarantee than those required in this work. There are also several CS algorithms for the support recovery where the objective is to determine the indices of the non-zero elements of the signal but not their values [12], [13]. The support recovery problem is indeed equivalent to the QGT problem. Notwithstanding, the existing schemes for support recovery rely on non-binary measurement matrices, and hence are not suitable for the QGT problem.

Last but not least, to the best of our knowledge, the majority of works on the CS problem focus mainly on the order optimality of the number of measurements, whereas in this work for the QGT problem we are also interested in minimizing the constant factor hidden in the order.

C. Main Contributions

In this work, we propose a non-adaptive quantitative group testing strategy for the sub-linear regime where $\frac{K}{N}$ vanishes as $K, N \to \infty$. We utilize sparse graph codes over bi-regular bipartite graphs and binary t-error-correcting BCH codes for the design of the proposed strategy. Leveraging powerful density evolution techniques for the analysis enables us not only to determine the exact value of constants in the number of tests needed but also to provide provable performance guarantees. We show that the proposed scheme provides exact recovery with probabilistic guarantee, i.e. recovers all the defective items with high probability. In particular, for the sub-linear regime, the proposed scheme requires at most $m \approx 1.19K \log_2\left(4.74\frac{N}{K}\right)$ tests to recover all the defective items with probability approaching one as $K, N \to \infty$. This bound can be achieved by $t = 2$. Moreover, for any $t \leq 4$, the testing and recovery algorithms of the proposed scheme have the computational complexity of $\mathcal{O}(K \log^2 \frac{N}{K})$ and $\mathcal{O}(K \log \frac{N}{K})$, respectively.

Due to the space constraints, the proofs of lemmas are not presented here, and can be found in [14].

II. Problem Setup and Notation

Throughout the paper, we use bold-face small and capital letters to denote vectors and matrices, respectively.

In this work, we consider the problem of quantitative group testing (QGT) with exact recovery and probabilistic guarantee, defined as follows. Consider a set of N items among which K items are defective. We focus on the sub-linear regime where the ratio $\frac{K}{N}$ vanishes as $K, N \to \infty$. The problem is to identify all the defective items with high probability while using minimum number of tests on subsets (groups) of the items, where the result of each test shows the number of defective items in the tested group.

Let $\mathbf{x} \in \{0,1\}^N$ represent the set of N items in which the coordinates with value 1 correspond to the defective items. A non-adaptive group testing problem consisting of m tests can be represented by a measurement matrix $\mathbf{A} \in \{0,1\}^{m \times N}$, where the i-th row of the matrix corresponds to the i-th test. That is, the coordinates with value 1 in the i-th row correspond to the items in the i-th test. The results of the m tests are expressed in the test vector $\mathbf{y} \in \{0, 1, \dots\}^m$, i.e.,

$$\mathbf{y} = [y_1, \cdots, y_m]^{\mathsf{T}} = \mathbf{A}\mathbf{x}. \tag{1}$$

The goal is to design a testing matrix $\mathbf{A}$ that has a small number of rows (tests), m, and can identify with high probability all the defective items given the test vector $\mathbf{y}$.

III. Proposed Scheme

A. Binary t-error-correcting codes and t-separable matrices

Definition 1. *(t-separable matrix) A binary matrix $\mathbf{D} \in \{0,1\}^{m \times n}$ (for $n > t$) is t-separable over field $\mathbb{F}$ if the sum (over field $\mathbb{F}$) of any set of t columns is distinct.*

By the definition, it can be easily seen that if a matrix $\mathbf{D}$ (with n columns) is t-separable over a field $\mathbb{F}$, then $\mathbf{D}$ is also s-separable over $\mathbb{F}$ for any $1 \leq s < t < n$.

The vector of test results, $\mathbf{y}$, is the sum of the columns in the testing matrix corresponding to the coordinates of the defective items. When a t-separable matrix over $\mathbb{R}$ is used as the testing matrix, the vector $\mathbf{y}$ will be distinct for any set of t defective items. Thus, a t-separable matrix over $\mathbb{R}$ can be used as the testing matrix for identifying t defective items. However, the construction of t-separable matrices for arbitrary t with minimum number of rows is an open problem. Instead, we can leverage the idea that the parity-check matrix of any binary t-error-correcting code is a t-separable matrix over $\mathbb{F}_2$. Note that t-separability over $\mathbb{F}_2$ results in t-separability over $\mathbb{R}$. Hence, a possible choice for designing a t-separable matrix over $\mathbb{R}$ is utilizing the parity-check matrix of a binary t-error-correcting code.

In this work, we use binary BCH codes for this purpose. It should be noted that Reed-Solomon (RS) codes can also be used; however, due to the restriction on testing matrix to be binary, RS codes are as efficient as BCH codes for the QGT problem. The key feature of the BCH codes which make them suitable for designing t-separable matrices is that it is possible to design binary BCH codes, capable of correcting any combination of t or fewer errors.

Definition 2. *[15] (Binary BCH codes) For any positive integers $m \geq 3$ and $t < 2^{m-1}$, there exists a binary t-error-correcting BCH code with the following parameters:*

$$\begin{cases} n = 2^m - 1 & \text{block length} \\ n - k \leq mt & \text{number of parity-check digits} \\ d_{\min} \geq 2t + 1 & \text{minimum Hamming distance} \end{cases}$$

The $t \times n$ parity-check matrix of such a code is given by $\mathbf{H}_t = \left(\alpha^{(2i-1)(j-1)}\right)_{i \in \{1,\cdots,t\}, j \in \{1,\cdots,n\}}$, where α is a primitive element in $\mathbb{F}_{2^m}$.

Since each entry of $\mathbf{H}_t$ is an element in $\mathbb{F}_{2^m}$, it can be represented by an m-tuple over $\mathbb{F}_2$. Thus, the number of rows in the binary representation of $\mathbf{H}_t$ is given by

$$R = tm = t\log_2(n+1). \tag{2}$$

B. Testing Algorithm

The design of the measurement matrix $\mathbf{A}$ in our scheme is based on an architectural philosophy that was proposed in [2] and [16]. The key idea is to design $\mathbf{A}$ using a sparse bi-regular bipartite graph and to apply a peeling-based iterative algorithm for recovering the defective items given $\mathbf{y}$.

Let $G_{\ell,r}(N,M)$ be a randomly generated bipartite graph where each of the N left nodes is connected to ℓ right nodes uniformly at random, and each of the M right nodes is connected to r left nodes uniformly at random. Note that there are $N\ell$ edge connections from the left side and Mr edge connections from the right side,

$$N\ell = Mr \tag{3}$$

Let $\mathbf{T}_G \in \{0,1\}^{M\times N}$ be the adjacency matrix of the graph $G_{\ell,r}(N,M)$, where each column in $\mathbf{T}_{\mathcal{G}}$ corresponds to a left node and has exactly ℓ ones, and each row corresponds to a right node and has exactly r ones. Let $\mathbf{t}_i \in \{0,1\}^N$ denote the i-th row of $\mathbf{T}_G$, i.e., $\mathbf{T}_G = [\mathbf{t}_1^\mathsf{T}, \mathbf{t}_2^\mathsf{T}, \cdots, \mathbf{t}_M^\mathsf{T}]^\mathsf{T}$. We assign s tests to each right node based on a signature matrix $\mathbf{U} \in \{0,1\}^{s\times r}$. The signature matrix is chosen as $\mathbf{U} = [\mathbf{1}_{1\times r}^\mathsf{T}, \mathbf{H}_t^\mathsf{T}]^\mathsf{T}$, where $\mathbf{1}_{1\times r}$ is an all-ones row of length r, and $\mathbf{H}_t \in \{0,1\}^{t\log_2(r+1)\times r}$ is the parity-check matrix of a binary t-error-correcting BCH code. The first row of the signature matrix counts the number of defective items connected to a right node, and given that this quantity is no more than t, the rest of the rows find the indices of the connected defective items. From (2), it can be easily seen that $s = R+1 = t\log_2(r+1)+1$.

The measurement matrix is given by $\mathbf{A} = [\mathbf{A}_1^\mathsf{T}, \cdots, \mathbf{A}_M^\mathsf{T}]^\mathsf{T}$ where $\mathbf{A}_i \in \{0,1\}^{s\times N}$ is a matrix that defines the s tests at the i-th right node. There are exactly r ones in each row $\mathbf{t}_i$ of $\mathbf{T}_G$, and the signature matrix $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_r]$ has r columns. Note that $\mathbf{u}_i = [1, \mathbf{h}_i^\mathsf{T}]^\mathsf{T}$ is the i-th column of $\mathbf{U}$, where $\mathbf{h}_i$ is the i-th column of $\mathbf{H}_t$. $\mathbf{A}_i$ is obtained by placing the r columns of $\mathbf{U}$ at the coordinates of the r ones of the row vector $\mathbf{t}_i$, and replacing zeros by all-zero columns,

$$\mathbf{A}_i = [\mathbf{0}, \ldots, \mathbf{0}, \mathbf{u}_1, \mathbf{0}, \ldots, \mathbf{u}_2, \mathbf{0}, \ldots, \mathbf{u}_r] \tag{4}$$

where $\mathbf{t}_i = [0, \ldots, 0,\ 1,\ 0, \ldots,\ 1,\ 0, \ldots,\ 1]$.

The number of rows in the measurement matrix $\mathbf{A}$, $m = M \times s$ where $s = t\log_2(r+1)+1$, represents the total number of tests in the proposed scheme.

C. Recovery algorithm

Let the observation vector corresponding to the i-th right node be defined as $\mathbf{z}_i = [z_{i,1}, z_{i,2}, \cdots, z_{i,s}]^\mathsf{T} = \mathbf{A}_i\mathbf{x}$, for all $\forall i \in \{1, \cdots, M\}$. Note that $\mathbf{z}_i = [y_{(i-1)s+1}, \cdots, y_{is}]^\mathsf{T}$.

Definition 3. *(t-*resolvable *right node) A right node is called* t-resolvable *if it is connected to t or fewer defective items.*

TABLE I: The function $c(t)$ and the optimal left degree ℓ.

t	1	2	3	4	5	6	7	8
$c(t)$	1.222	0.597	0.388	0.294	0.239	0.202	0.176	0.156
ℓ	3	2	2	2	2	2	2	2

The recovery algorithm performs in rounds as follows. In each round, the recovery algorithm first iterates through all the right node observation vectors $\{\mathbf{z}_i\}_{i=1}^M$, and resolves all the t-resolvable right nodes (for more details, see the proof of [14, Lemma 1]). Then, given the identities of the recovered left nodes, the edges connected to these defective items are peeled off the graph. That is, the contributions of the recovered defective items will be removed from the unresolved right nodes so that new right nodes may become t-resolvable for the next round. The recovery algorithm terminates when there is no more t-resolvable right nodes.

IV. Main Results

In this section, we present our main results. Theorem 1 characterizes the required number of tests that guarantees the identification of all the defective items with probability approaching one as $K, N \to \infty$. Theorem 2 presents the computational complexity of the proposed scheme. The proofs of Theorems 1 and 2 are given in Section V.

Theorem 1. *For the sub-linear regime, the proposed scheme recovers all defective items with probability approaching one (as $K, N \to \infty$) with at most $m = c(t)K\left(t\log_2\left(\frac{\ell N}{c(t)K}+1\right)+1\right)+1$ tests, where $c(t)$ depends only on t. Table I shows the values of $c(t)$ for $t \le 8$.*

Theorem 2. *The testing and recovery algorithms of the proposed scheme for any $t \le 4$ have the computational complexity of $\mathcal{O}(K\log^2\frac{N}{K})$ and $\mathcal{O}(K\log\frac{N}{K})$, respectively.*

V. Proofs of Main Theorems

A. Proof of Theorem 1

Consider a set of N items out of which K items are defective. Note that in the QGT problem, performing one initial test (on all items) would suffice to obtain the number of defective items. As mentioned in Section III-C, our scheme employs an iterative recovery algorithm. In each iteration, the algorithm finds and resolves all the *t-resolvable* right nodes. At the end of each iteration, the decoder subtracts the contribution of the identified defective items from the unresolved right nodes. This process is repeated until there is no *t-resolvable* right nodes left in the graph. The fraction of defective items that remain unidentified at the end of each iteration can be analyzed using density evolution as follows.

Assuming that the exact number of the defective items, K, is known and the values assigned to the defective and non-defective items are one and zero, respectively, the left-and-right-regular bipartite graph can be pruned. All the zero left nodes and their respective edges are removed from the graph. The number of left nodes in the pruned graph is K, but the degree of these nodes remains unchanged. On the other hand, the number of right nodes remains unchanged, but the resulting graph is not right-regular any longer.

Let λ be the average right degree, i.e., $\lambda = \frac{K\ell}{M}$. Let $\rho(x) \triangleq \sum_{i=1}^{\min(K,r)} \rho_i x^{i-1}$ be the right edge degree distribution, where ρ_i is the probability that a randomly picked edge in the pruned graph is connected to a right node of degree i, and $\min(K,r)$ is the maximum degree of a right node. As shown in [16], as $K, N \to \infty$, we have $\rho_i = e^{-\lambda}\frac{\lambda^{i-1}}{(i-1)!}$.

Lemma 1. *Let p_j be the probability that a randomly chosen defective item is not recovered at iteration j of the recovery algorithm; and let q_j be the probability that a randomly picked right node is resolved at iteration j of the recovery algorithm. The relation between p_j and p_{j+1} is determined by the following density evolution equations:*

$$q_j = \sum_{i=1}^{t} \rho_i + \sum_{i=t+1}^{\min(K,r)} \rho_i \sum_{k=0}^{t-1} \binom{i-1}{k} p_j^k (1-p_j)^{i-k-1}, \quad (5)$$

$$p_{j+1} = (1-q_j)^{\ell-1}, \quad (6)$$

where t is the level of separability, and ρ_i is the probability that a randomly picked edge in the pruned graph is connected to a right node of degree i.

Note that p_j is only a function of the variables t, ℓ, and λ when $\min(K,r) \to \infty$. Recall that the goal is to minimize the total number of tests $m = M \times s$, where M is the number of right nodes, and s is the number of rows in the signature matrix. The number of rows, s, in the signature matrix depends only on the level of separability, t. For a given t, we can minimize the number of right nodes $M = \frac{\ell}{\lambda}K$ subject to the constraint $\lim_{j\to\infty} p_j(\ell,\lambda) = 0$, so as to minimize the total number of the tests. The constraint $\lim_{j\to\infty} p_j(\ell,\lambda) = 0$ guarantees that running the recovery algorithm for sufficiently large number of iterations, the probability that a randomly chosen defective item remains unidentified approaches zero. For any $\ell \geq 2$, let $\lambda_T(\ell) \triangleq \sup\{\lambda : \lim_{j\to\infty} p_j(\ell,\lambda) = 0\}$. Then, for any $\ell \geq 2$ and $\lambda < \lambda_T(\ell)$, we have $\lim_{j\to\infty} p_j(\ell,\lambda) = 0$. Accordingly, for any $\ell \geq 2$ and $M = \frac{\ell}{\lambda}K > \frac{\ell}{\lambda_T(\ell)}K$, it follows that $\lim_{j\to\infty} p_j(\ell,\lambda) = 0$. Our goal is then to compute

$$\min_{\ell\in\{2,3,\dots\}} \frac{\ell}{\lambda_T(\ell)}K. \quad (7)$$

We can solve this problem numerically and attain the optimal $\ell^\star$. Let $c(t) \triangleq \frac{\ell^\star}{\lambda_T(\ell^\star)}$. The number of right nodes can be chosen as $M = c(t)K\beta$ for any $\beta > 1$ to guarantee that $M > c(t)K = \frac{\ell^\star}{\lambda_T(\ell^\star)}K$. Substituting $M = c(t)K\beta$ in (3) results in $r = \frac{\ell N}{c(t)K\beta}$. Therefore, the total number of tests will become $m = M \times s = c(t)K\beta\left(t\log_2\left(\frac{\ell N}{c(t)K\beta}+1\right)+1\right)$.

Lemma 2. *There exist some $\beta > 1$ such that*

$$c(t)K\left(t\log_2\left(\frac{\ell N}{c(t)K}+1\right)+1\right)+1 \geq c(t)K\beta\left(t\log_2\left(\frac{\ell N}{c(t)K\beta}+1\right)+1\right).$$

From Lemma 2, it follows that with probability approaching one as $K, N \to \infty$, $c(t)K\left(t\log_2\left(\frac{\ell N}{c(t)K}+1\right)+1\right)+1$ tests would suffice for the proposed scheme to recover all the defective items. This completes the proof.

B. Proof of Theorem 2

Lemma 3. *For any $t \leq 4$, the computational complexity of resolving each t-resolvable right node is $\mathcal{O}(\log r)$.*

This can be done by first finding the error locator polynomial using the Berlekamp-Massey algorithm [15] with complexity $\mathcal{O}(t^2 \log r)$, and then finding the roots of this polynomial directly using the algorithm of [17] with complexity $\mathcal{O}(t \log r)$.

The total number of right nodes, M, is $\mathcal{O}(K)$. From Lemma 3, it then follows that the complexity of the recovery algorithm is $\mathcal{O}(K \log r)$. Using (3), it is easy to see that for any $t \leq 4$ the recovery algorithm has complexity $\mathcal{O}(K \log \frac{N}{K})$. The total number of measurements is m and for each measurement r summations are performed. Hence, the complexity of the testing algorithm is $\mathcal{O}(mr)$, which becomes equivalent to $\mathcal{O}(K \log^2 \frac{N}{K})$ for any $t \leq 4$.

VI. Evaluation of $c(t)$

In this section, we present the complete analysis for the case of $t = 1$, and show how one can evaluate $c(t)$ at $t = 1$, i.e., $c(1)$. The same procedure can be used for evaluating $c(t)$ at any $t > 1$.

For the case of $t = 1$, the density evolution equations (5) and (6) can be combined as $p_{j+1} = (1 - \sum_{i=1}^{\min(K,r)} \rho_i (1-p_j)^{i-1})^{\ell-1}$. Clearly, $p_1 = 1$. Substituting $\rho_i = e^{-\lambda}\frac{\lambda^{i-1}}{(i-1)!}$, we can rewrite this equation as $p_{j+1} = (1 - e^{-\lambda}\sum_{i=1}^{\min(K,r)} \frac{\lambda^{i-1}}{(i-1)!}(1-p_j)^{i-1})^{\ell-1}$. Letting $\min(K,r) \to \infty$, we get $p_{j+1} = \left(1 - e^{-\lambda p_j}\right)^{\ell-1}$.

Lemma 4. *For any $\ell \geq 2$ and any $\lambda > 0$, the infinite sequence $\{p_1, p_2, \cdots\}$ converges.*

Lemma 5. *Let p_∞ be the limit of the sequence $\{p_1, p_2, \cdots\}$, and let*

$$\lambda_T \triangleq \inf_{0<x<1}\left(\frac{\ln(1 - x^{\frac{1}{\ell-1}})}{-x}\right).$$

Then, for any $\ell \geq 2$, we have

$$\begin{cases} p_\infty = 0, & 0 < \lambda < \lambda_T, \\ p_\infty > 0, & \lambda \geq \lambda_T. \end{cases}$$

To solve the optimization problem in (7), for any $\ell \geq 2$ the corresponding λ_T must be computed. Then, the optimal $\ell^\star$ and $c(1) = \frac{\ell^\star}{\lambda_T(\ell^\star)}$ will be attained easily.

VII. Comparison Results

In this section we will evaluate the performance of the proposed scheme based on our theoretical analysis and the Monte Carlo simulations.

Based on the results in Theorem 1 and Table I, Fig. 1 depicts the total number of tests (m) required to identify all K defective items among a total of $N = 2^{16}$ items,

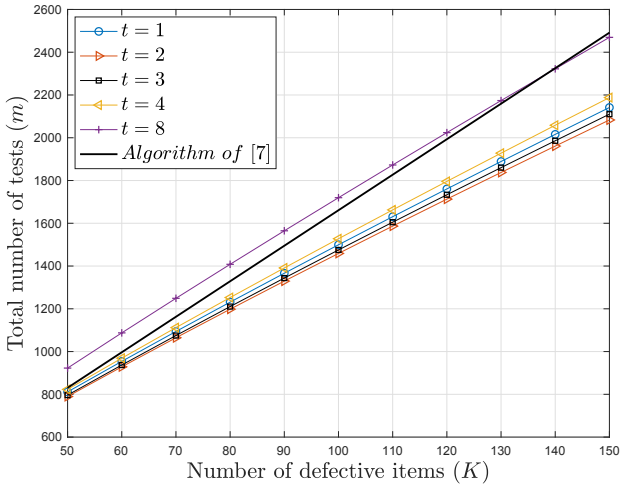

Fig. 1: The number of required tests (m) to identify all K defective items (for different values of K) among $N = 2^{16}$ items for different values of t obtained via analysis.

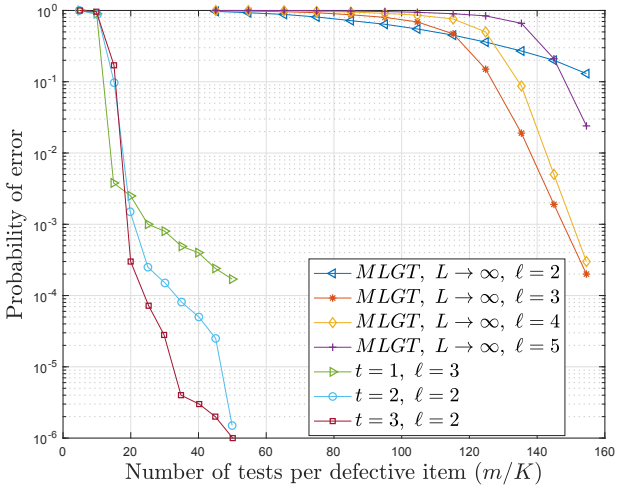

Fig. 2: The average fraction of unidentified defective items obtained via Monte Carlo simulations for $N = 2^{16}$ items among which $K = 100$ items are defective.

for different values of t. For reference, the performance of the best known polynomial-time ($O(KN \log N)$) algorithm proposed in [7] is also shown in Fig. 1. As it can be seen, when $t \in \{1, 2, 3\}$ the required number of tests for identifying all the defective items is less than that for larger values of t. The results of our theoretical analysis reveal that the minimum number of required tests for the proposed scheme is achieved by $t = 2$.

Using the Monte Carlo simulation, we also compare the performance of the proposed scheme for $t \in \{1, 2, 3\}$ with the performance of the Multi-Level Group Testing (MLGT) scheme from [9]. The MLGT scheme is a semi-quantitative group testing scheme where the result of each test is an integer in the set $\{0, 1, 2, \cdots, L\}$. Letting $L \to \infty$, the MLGT scheme becomes a QGT scheme. Based on the optimization that we have performed, the optimal left degree for the MLGT scheme is $\ell^\star = 3$ when $L \to \infty$. For $K = 100$ defective items among a population of $N = 2^{16}$ items, the probability of error, defined as the probability of a defective item to remain unidentified, for the MLGT scheme and the proposed scheme are shown in Fig. 2 for different values of m/K. As it can be observed, the proposed scheme for all the three tested values of t outperforms the MLGT scheme significantly. For instance, when the probability of error is 2×10^{-4}, the required number of tests for the MLGT scheme (for $\ell = 3$) is 3, 5, and 7 times more than that of the proposed scheme for $t = 1$, 2, and 3, respectively.

Simulation results suggest a trade-off between the total number of tests and the probability of error as functions of the variable t. For larger values of t, the proposed scheme requires more number of tests whereas the probability of error tends to zero faster.

REFERENCES

[1] J. Scarlett and V. Cevher, "How little does non-exact recovery help in group testing?" in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017, pp. 6090–6094.

[2] K. Lee, R. Pedarsani, and K. Ramchandran, "SAFFRON: A fast, efficient, and robust framework for group testing based on sparse-graph codes," *CoRR*, vol. abs/1508.04485, 2015.

[3] C. Wang, Q. Zhao, and C. N. Chuah, "Optimal nested test plan for combinatorial quantitative group testing," *IEEE Transactions on Signal Processing*, vol. PP, no. 99, 2017.

[4] H. S. Shapiro, "Problem E 1399," *Amer. Math. Monthly*, vol. 67, no. 82, pp. 697–697, 1960.

[5] N. H. Bshouty, "Optimal algorithms for the coin weighing problem with a spring scale," in *Conference on Learning Theory*, 2009.

[6] E. Karimi, F. Kazemi, A. Heidarzadeh, and A. Sprintson, "A simple and efficient strategy for the coin weighing problem with a spring scale," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 1730–1734.

[7] B. Lindström, "Determining subsets by unramified experiments," in *A survey of Statistical Design and Linear Models*, 1975.

[8] V. Grebinski and G. Kucherov, "Optimal reconstruction of graphs under the additive model," *Algorithmica*, vol. 28, no. 1, pp. 104–124, 2000.

[9] P. Abdalla, A. Reisizadeh, and R. Pedarsani, "Multilevel group testing via sparse-graph codes," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*, Oct 2017, pp. 895–899.

[10] M. Lotfi and M. Vidyasagar, "A fast noniterative algorithm for compressive sensing using binary measurement matrices," *IEEE Trans. on Signal Processing*, vol. 66, no. 15, pp. 4079–4089, Aug 2018.

[11] M. Iwen, "Compressed sensing with sparse binary matrices: Instance optimal error guarantees in near-optimal time," *Journal of Complexity*, vol. 30, no. 1, pp. 1–15, 2014.

[12] X. Li, S. Pawar, and K. Ramchandran, "Sub-linear time support recovery for compressed sensing using sparse-graph codes," *CoRR*, vol. abs/1412.7646, 2018. [Online]. Available: http://arxiv.org/abs/1412.7646

[13] J. Haupt and R. Baraniuk, "Robust support recovery using sparse compressive sensing matrices," in *2011 45th Annual Conference on Information Sciences and Systems*, March 2011, pp. 1–6.

[14] E. Karimi, F. Kazemi, A. Heidarzadeh, K. R. Narayanan, and A. Sprintson, "Sparse graph codes for non-adaptive quantitative group testing," *CoRR*, vol. abs/1901.07635, 2019. [Online]. Available: http://arxiv.org/abs/1901.07635

[15] S. Lin and D. J. Costello, *Error control coding*. Pearson Education India, 2001.

[16] A. Vem, N. T. Janakiraman, and K. R. Narayanan, "Group testing using left-and-right-regular sparse-graph codes," *CoRR*, vol. abs/1701.07477, 2017. [Online]. Available: http://arxiv.org/abs/1701.07477

[17] C.-L. Chen, "Formulas for the solutions of quadratic equations over GF(2^m)(corresp.)," *IEEE Transactions on Information Theory*, vol. 28, no. 5, pp. 792–794, 1982.