Graph-Constrained Group Testing

Mahdi Cheraghchi, Amin Karbasi and Soheil Mohajer
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland

Venkatesh Saligrama
Department of Electrical and Computer Engineering
Boston University
MA-02215, USA

***Abstract*—Non-adaptive group testing involves grouping arbitrary subsets of n items into different pools and identifying defective items based on tests obtained for each pool. Motivated by applications in network tomography, sensor networks and infection propagation we formulate non-adaptive group testing problems on graphs. Unlike conventional group testing problems each group here must conform to the constraints imposed by a graph. For instance, items can be associated with vertices and each pool is any set of nodes that must be path connected. In this paper we associate a test with a random walk. In this context conventional group testing corresponds to the special case of a complete graph on n vertices.**

For interesting classes of graphs we arrive at a rather surprising result, namely, that the number of tests required to identify d defective items is substantially similar to that required in conventional group testing problems, where no such constraints on pooling is imposed. Specifically, if $T(n)$ corresponds to the mixing time of the graph G, we show that with $m = O(d^2T^2(n)\log(n/d))$ non-adaptive tests, one can identify the defective items. Consequently, for the Erdős-Rényi random graph $G(n,p)$, as well as expander graphs with constant spectral gap, it follows that $m = O(d^2\log^3 n)$ non-adaptive tests are sufficient to identify d defective items. We next consider a specific scenario that arises in network tomography and show that $m = O(d^3\log^3 n)$ non-adaptive tests are sufficient to identify d defective items. We also consider noisy counterparts of the graph constrained group testing problem and develop parallel results for these cases.

I. Introduction

In this paper we introduce the graph constrained group testing problem (GCGT) motivated by applications in network tomography, sensor networks and infection propagation. While group testing theory (see [1], [2] and more recently [3]), and its numerous applications, such as industrial quality assurance [4], DNA library screening [5], software testing [6], and multi-access communications [7], have been systematically explored, the graph constrained group testing problem is new to the best of our knowledge.

Group testing involves identifying at most d defective items out of a set of n items. In non-adaptive group testing, which is the subject of this paper, we are given an $m \times n$ binary matrix, M, usually referred to as a test or measurement matrix. Ones on the jth row of M indicate which subset of the n items belongs to the jth pool. A test is conducted on each pool; a positive outcome indicating that a defective item is part of the pool; and a negative test indicating that no defective items are part of the pool. The conventional group testing problem is to design a matrix M with minimum number of rows m that guarantees error free identification of the defective items. While the best known (probabilistic) pooling design requires a test matrix with $m = O(d^2\log(n/d))$ rows, and an almost-matching lower bound of $m = \Omega(d^2(\log n)/(\log d))$ is known on the number of pools (cf. [2, Chapter 7]), the size of the optimal test still remains open.

Note that in the standard group testing problem the test matrix M can be designed arbitrarily. In this paper we consider a generalization of the group testing problem to the case where the matrix M must conform to constraints imposed by a graph $G = (V, E)$. In general the problem we describe naturally arises in several applications such as network tomography [8], [9], sensor networks [10], and infection propagation [11]. While the graph constrained group testing problem has been alluded to in these applications the problem of test design or the characterization of the minimum number of tests, to the best of our knowledge, has not been addressed before. In this light our paper is the first to formalize the graph constrained group testing (GCGT) problem[1]. In the GCGT problem the n items are either vertices or links (edges) of the graph; at most d of them are defective. The task is to identify the defective vertices or edges. The test matrix M is constrained as follows: for items associated with vertices each row must correspond to a subset of vertices that are connected by a path on the graph; for items associated with links each row must correspond to links that are path connected in the line graph of G. The task is to design an $m \times n$ binary test matrix with minimum number of rows m that guarantees error free identification of the defective items.

The GCGT problem has close connections to network tomography [9], [8], [10], which deals with identification of congested links from end-to-end path measurements for a given network. Congested links lead to packet losses or delays for a route that goes through these links. A network is associated with a graph $G = (V, E)$ where the set V denotes the network routers/hosts and the set E denotes

Emails: {mahdi.cheraghchi, amin.karbasi, soheil.mohajer}@epfl.ch, and srv@bu.edu. M.C. is supported by the ERC Advanced Investigator Grant 228021 of A. Shokrollahi. V.S. is supported by the U.S. Department of Homeland Security under Award Number 2008-ST-061-ED0001 and NSF CAREER Award Number ECS 0449194.

[1]While variations of classical group testing possessing a graph theoretic nature has been studied, our setting where pools are associated with paths is new. Notable examples of employing graph constraints include the problem of learning hypergraphs [12]. Another variation concerns group testing with constraints defined by a rooted tree (see [2, Chapter 12]).

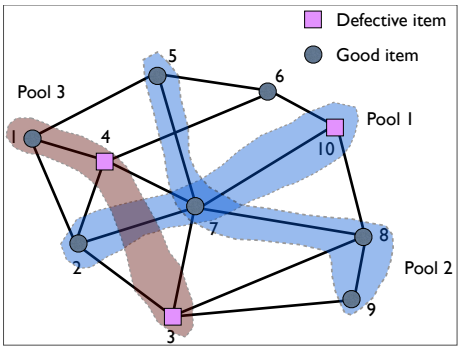

Fig. 1. The result of pool 1 is positive since it contains a defective item, whereas the result of pool 2 is negative since it does not contain a defective item. Pool 3 is not consistent with the graph and thus not allowed since the items are not connected by a path.

the communication links (see Fig. 1). Suppose, we have a monitoring system that consists of one or more end hosts (so called *vantage* points) that can send and receive packets. Each vantage point sends packets through the network by assigning the routes and the end hosts. A vantage point can only assign those routes which form a path in the graph G. Each measurement has a binary outcome, namely, it is deemed to be positive if packets sent along a route have significant path losses or path delays (and negative otherwise). The question of interest is to determine the number of measurements that is needed in order to identify the congested links in a given network.

Nevertheless, the network tomography problem just described is static [9] in that it corresponds to the case where the binary test matrix M is fixed once a set of vantage points has been chosen. Our paper is closely related to situations arising in wireless sensor networks (WSN), where the test matrices can be random. Indeed in WSNs [10] the routing topology is constantly changing. At any instant, sensor nodes form a tree to route packets to the sink. The routing tree constantly changes unpredictably but must be consistent with the underlying network connectivity.

Motivated by the WSN scenario we describe pool designs based on random walks on graphs. As is well known a random walk is the state evolution on a finite reversible Markov chain. Each row of the binary test matrix is derived from the evolution of the random walk, namely, the ones on the jth row of M correspond to the vertices visited by the jth walk. This is close to the WSN scenario because as in the WSN scenario the path between two given nodes changes randomly. However, a significant difference is that random walks have loops. In this context it is worth mentioning that there is a close connection between random walks and random spanning trees [13].

We first consider random walks that start either at a random node or an arbitrary node but terminate in some appropriately chosen time t. By optimizing the length of the walk we arrive at a surprising result for interesting classes of graphs. Specifically we show that the number of tests required to identify d defective items is substantially similar to that required in conventional group testing problems, where no such constraints on pooling is imposed. The best known result for the number of tests required when no graphical constraints are imposed scales as $O(d^2 \log(n/d))$. For the graph constrained case we show that with $m = O(d^2 T^2(n) \log(n/d))$ non-adaptive tests one can identify the defective items, where $T(n)$ corresponds to the mixing time of the graph G. Consequently, for the Erdős-Rényi random graph $G(n,p)$, as well as expander graphs with constant spectral gap, it follows that $m = O(d^2 \log^3 n)$ non-adaptive tests are sufficient to identify d defective items. Next we consider unbounded-length random walks that originate at a source node and terminate at a sink node. Both the source node and the sink node can either be arbitrary or be chosen uniformly at random. This situation is closer in spirit to the WSN network tomography problem. In this scenario we show that $m = O(d^3 \log^3 n)$ non-adaptive tests are sufficient to identify d defective items.

Finally, we also consider noisy counterparts of the graph constrained group testing problem and develop parallel results for these cases. Specifically, we consider the so called *dilution model*. In this case each item can be *diluted* in each test with some a priori known probability. This corresponds to the case when a test on a path with a congested link can turn out to be negative with some probability. We show that similar scaling results holds for this case as well.

The rest of this paper is organized as follows. In Section II, we introduce our notation and mention some basic facts related to group testing and random walks on graphs. Then Section III formally describes the problem that we consider and states our main results. The reader is referred to the full version of this paper[2] for an elaborate discussion on various applications of graph-constrained group testing and omitted proof details.

II. Definitions and Notation

In this section we introduce some tools, definition and notations which are used throughout the paper.

Definition 1: For two given boolean vectors S and T of the same length we denote their element-wise logical or by $S \vee T$. More generally, we will use $\bigvee_{i=1}^{d} S_i$ to denote the element-wise or of d boolean vectors. The logical subtraction of two boolean vectors $S = (s_1, \ldots, s_n)$ and $T = (t_1, \ldots, t_n)$, denoted by $S \setminus T$, is defined as a boolean vector which has a 1 at each position i if and only if $s_i = 1$ and $t_i = 0$. We also use $|S|$ to show the number of 1's in (i.e., the Hamming weight of) a vector S.

Matrices that are suitable for the purpose of group testing are known as *disjunct* matrices. The formal definition is as follows:

Definition 2: An $m \times n$ boolean matrix M is called d-disjunct, if, for every column S_0 and every choice of d columns $S_1, \ldots, S_d$ of M (different from S_0), there is at least one row at which the entry corresponding to S_0 is 1 and those corresponding to $S_1, \ldots, S_d$ are all zeros. More generally, for an integer $e \geq 0$, the matrix is called (d,e)-disjunct if for

[2]Available online at ⟨http://arxiv.org/abs/1001.1445⟩

every choice of the columns S_i as above, they satisfy

$$|S_0 \setminus \bigvee_{i=1}^{d} S_i| > e.$$

A matrix that is $(d, 0)$-disjunct is said to be d-disjunct.

A classical observation in group testing theory states that disjunct matrices can be used in non-adaptive group testing schemes to distinguish sparse boolean vectors (cf. [2]). More precisely, suppose that a d-disjunct matrix M with n columns is used as the measurement matrix; i.e., we assume that the rows of M are the characteristic vectors of the pools defined by the scheme. Then, the test outcomes obtained by applying the scheme on two distinct d-sparse vectors of length n must differ in at least one position. More generally, if M is taken to be (d, e)-disjunct, the test outcomes must differ in at least $e+1$ positions. Thus, the more general notion of (d, e)-disjunct matrices is useful for various "noisy" settings, where we are allowed to have a few false outcomes. As observed in [11], this notion is also suitable for handling the *dilution* model, where with some small probability a defective item may not affect the outcome of a test in which it participates.

For our application, sparse vectors correspond to boolean vectors encoding the set of defective vertices, or edges, in a given undirected grph. Moreover, we aim to construct measurement matrices that are not only disjunct in the classical sense describe above, but are also constrained to be *consistent* with the underlying graph, as formalized below.

Definition 3: Let $G = (V, E)$ be an undirected graph, and A and B be boolean matrices with $|V|$ and $|E|$ columns, respectively. The columns of A are indexed by the elements of V and the columns of B are indexed by the elements of E. Then,

- The matrix A is said to be *vertex-consistent* with G if each row of A, seen as the characteristic vector of a subset of V, exactly represents the set of vertices visited by some walk on G.
- The matrix B is said to be *edge-consistent* with G if each row of B, seen as the characteristic vector of a subset of E, exactly corresponds to the set of edges traversed by a walk on G.

Note that the choice of the walk corresponding to each row of A or B need not be unique. Moreover, a walk may visit a vertex (or edge) more than once.

Definition 4: An undirected graph $G = (V, E)$ is called (D, c)-uniform, for some $c \geq 1$, if the degree of each vertex $v \in V$ (denoted by $\deg(v)$) is between D and cD.

Throughout this work, the constraint graphs are considered to be (D, c)-uniform, for an appropriate choice of D and some (typically constant) parameter c. When $c = 1$, the graph is D-regular.

Definition 5: The *point-wise distance* of two probability distributions μ, μ' on a finite space Ω is defined as

$$\|\mu - \mu'\|_\infty := \max_{i \in \Omega} |\mu(i) - \mu'(i)|,$$

where $\mu(i)$ (resp., $\mu'(i)$) denote the probability assigned by μ (resp., μ') to the outcome $i \in \Omega$.

Definition 6: Let $G = (V, E)$ with $|V| = n$ be a (D, c)-uniform graph and denote by μ its stationary distribution. The *δ-mixing time* of G (with respect to the ℓ_∞ norm) is the smallest integer t such that a random walk of length t starting at any vertex in G ends up having a distribution μ' with $\|\mu' - \mu\|_\infty \leq \delta$. For concreteness, we define the quantity $T(n)$ as the δ-mixing time of G for $\delta := (1/2cn)^2$.

Two classes of graphs that are of particular importance for us are expander graphs and Erdős-Rényi random graphs.

Let G be a regular undirected graph[3], and denote by A the normalized adjacency matrix of G (so that the entries on each row/column of A sum up to 1). The *spectral gap* of G is defined as $1 - \lambda$, where λ is the second largest eigenvalue of A in absolute value. We consider G an *expander graph* when the spectral gap is constant; i.e., $1 - \lambda = \Omega(1)$.

An instance of a random graph in the Erdős-Rényi model is obtained using the following procedure: Start with the complete graph on n vertices, and remove each edge of the graph independently with probability $1 - p$. We denote the corresponding sample space by $G(n, p)$.

III. Problem setting and Main Results

Problem Statement. Consider a given graph $G = (V, E)$ in which at most d vertices (resp., edges) are defective. The goal is to characterize the set of defective items using as small number of measurements as possible, where each measurement determines whether the set of vertices (resp., edges) observed along a path on the graph has a non-empty intersection with the defective set. We call the problem of finding defective vertices *vertex group testing* and that of finding defective edges *edge group testing*.

As mentioned earlier, not all sets of vertices can be grouped together, and only those that share a path on the underlying graph G can participate in a pool (see Fig. 1).

In the following, we introduce four random constructions (designs) for both problems. The proposed designs follow the natural idea of determining pools by taking random walks on the graph.

Design 1. *Given:* a constraint graph $G = (V, E)$ with $r \geq 0$ designated vertices $s_1, \ldots, s_r \in V$, and integer parameters m and t.
Output: an $m \times |V|$ boolean matrix M.
Construction: Construct each row of M independently as follows: Let $v \in V$ be any of the designated vertices s_i, or otherwise a vertex chosen uniformly at random from V. Perform a random walk of length t starting from v, and let the corresponding row of M be the characteristic vector of the set of vertices visited by the walk.

By construction, Designs 1 and 3 (resp., Designs 2 and 4) output boolean matrices that are vertex- (resp., edge-)

[3]To be precise, we have implicitly assumed that G belongs to an infinite family of regular graphs containing arbitrarily large graphs.

Design 2.
Given: a constraint graph $G = (V, E)$ and integer parameters m and t.
Output: an $m \times |E|$ boolean matrix M.
Construction: Construct each row of M independently as follows: Let $v \in V$ be any arbitrary vertex of G. Perform a random walk of length t starting from v, and let the corresponding row of M be the characteristic vector of the set of edges visited by the walk.

Design 3.
Given: a constraint graph $G = (V, E)$ with $r \geq 0$ designated vertices $s_1, \ldots, s_r \in V$, a *sink node* $u \in V$, and integer parameter m.
Output: an $m \times |V|$ boolean matrix M.
Constructions: Construct each row of M independently as follows: Let $v \in V$ be any of the designated vertices s_i, or otherwise a vertex chosen uniformly at random from V. Perform a random walk starting from v until we reach u, and let the corresponding row of M be the characteristic vector of the set of vertices visited by the walk.

Design 4.
Given: a constraint graph $G = (V, E)$, a *sink node* $u \in V$, and integer parameter m.
Output: an $m \times |E|$ boolean matrix M.
Construction: Construct each row of M independently as follows: Let $v \in V$ be any arbitrary vertex of G. Perform a random walk, starting from v until we reach u, and let the corresponding row of M be the characteristic vector of the set of edges visited by the walk.

consistent with the graph G. Our main goal is to show that, when the number of rows m is sufficiently large, the output matrices become d-disjunct (for a given parameter d) with overwhelming probability.

Remark 7: Designs 1 and 3 in particular provide two choices for constructing the measurement matrix M. Namely, the start vertices can be chosen within a fixed set of designated vertices, or, chosen randomly among all vertices of the graph. As we will see later, in theory there is no significant difference between the two schemes. However, for some applications it might be the case that only a small subset of vertices are accessible as the starting points (e.g., in network tomography such a subset can be determined by the vantage points), and this can be modeled by an appropriate choice of the designated vertices in Designs 1 and 3.

The following theorem states the main result of this work.

Theorem 8: Let $p \geq 0$ be a fixed parameter, and suppose that $G = (V, E)$ is a (D, c)-uniform graph on n vertices with mixing time $T(n)$. Then there exist parameters with asymptotic values given in Table I such that, provided that $D \geq D_0$,

1) Design 1 with the path length $t := t_1$ and the number of measurements $m := m_1$ outputs a matrix M that is vertex-consistent with G. Moreover, once the columns of M corresponding to the designated vertices $s_1, \ldots, s_r$ are removed, the matrix becomes d-disjunct with probability $1 - o(1)$. More generally, for $m := m'_1$ the matrix becomes (d, e_1)-disjunct with probability $1 - o(1)$.
2) Design 2 with path length $t := t_2$ and $m := m_2$ measurements outputs a matrix M that is edge-consistent with G and is d-disjunct with probability $1 - o(1)$. More generally, for $m := m'_2$ the matrix becomes (d, e_2)-disjunct with probability $1 - o(1)$.
3) Design 3 with the number of measurements $m := m_3$ outputs a matrix M that is vertex-consistent with G. Moreover, once the columns of M corresponding to the designated vertices $s_1, \ldots, s_r$ and the sink node u are removed, the matrix becomes d-disjunct with probability $1 - o(1)$. More generally, for $m := m'_3$ the matrix becomes (d, e_3)-disjunct with probability $1 - o(1)$.
4) Design 4 with the number of measurements $m := m_4$ outputs a matrix M that is edge-consistent with G and is d-disjunct with probability $1 - o(1)$. More generally, for $m := m'_4$ the matrix becomes (d, e_4)-disjunct with probability $1 - o(1)$.

TABLE I
THE ASYMPTOTIC VALUES OF VARIOUS PARAMETERS IN THEOREM 8.

Parameter	Value
D_0	$O(c^2 dT^2(n))$
m_1, m_2	$O(c^4 d^2 T^2(n) \log(n/d))$
m_3	$O(c^8 d^3 T^4(n) \log(n/d))$
m_4	$O(c^9 d^3 DT^4(n) \log(n/d))$
t_1	$O(n/(c^3 dT(n)))$
t_2	$O(nD/(c^3 dT(n)))$
e_1, e_2, e_3, e_4	$\Omega(pd \log(n/d)/(1-p)^2)$
$m'_i, i \in [4]$	$O(m_i/(1-p)^2)$

Proof (sketch): First, observe that by construction, the obtained matrices are consistent with the underlying graph. Let W_1 be any of the m walks performed in Design 1, and similarly, W_2, W_3, W_4 be any of the walks in Designs 2, 3, and 4, respectively. We distinguish the following quantities related to the walks:

- For a vertex $v \in V$ and subset $A \subseteq V$ such that $v \notin A$, $|A| \leq d$, denote by $\pi_{v,A}$ (resp., $\pi^{(u)}_{v,A}$) the probability that W_1 (resp., W_3) passes v but none of the vertices in A. We assume that $\{v\} \cup A$ is disjoint from $\{s_1, \ldots, s_r\}$ in Design 1 and from $\{s_1, \ldots, s_r, u\}$ in Design 3.
- For an edge $e \in E$ and subset $B \subseteq E$ such that $e \notin B$, $|B| \leq d$, denote by $\pi_{e,B}$ (resp., $\pi^{(u)}_{e,B}$) the probability that W_2 (resp., W_4) passes e but none of the edges in B.

Under the assumptions on the underlying graph G, it can be shown that,

$$\pi_{v,A} = \Omega\left(\tfrac{1}{c^4 dT^2(n)}\right) \quad \pi_{e,B} = \Omega\left(\tfrac{1}{c^4 dT^2(n)}\right)$$
$$\pi^{(u)}_{v,A} = \Omega\left(\tfrac{1}{c^8 d^2 T^4(n)}\right) \quad \pi^{(u)}_{e,B} = \Omega\left(\tfrac{1}{c^9 d^2 DT^4(n)}\right).$$

The proofs for the above lower bounds are technical and rather lengthy, and are skipped in this short presentation. Detailed proofs can be found in the full version of this paper. For concreteness, we focus on the first claim about Design 1.

The proof for the other designs is similar. Take a vertex $v \in V$ and $A \subseteq V$ such that $v \notin A$, $|A| \leq d$, and $(\{v\} \cup A) \cap \{s_1, \ldots, s_r\} = \emptyset$. For each $i \in [m_1]$, define a random variable $X_i \in \{0,1\}$ such that $X_i = 1$ iff the ith row of M has a 1 entry at the column corresponding to v and zeros at those corresponding to the elements of A. Let $X := \sum_{i=1}^{m_1} X_i$. Note that the columns corresponding to v and A violate the disjunctness property of M iff $X = 0$, and that the X_i are independent Bernoulli random variables. For each i, $X_i = 1$ happens exactly when the ith random walk passes the vertex v but never hits any vertex in A, and by the lower bound on $\pi_{v,A}$, we have $\Pr[X_i = 1] = \Omega(1/(c^4 dT^2(n)))$.

Denote by p_f the *failure probability*, namely that the resulting matrix M is not d-disjunct. By a union bound on the choice of v and A, we get

$$\begin{aligned} p_f &\leq \sum_{v,A} (1 - \pi_{v,A})^{m_1} \\ &\leq \exp\left(d \log \frac{n}{d}\right) \cdot \left(1 - \Omega\left(\frac{1}{c^4 dT^2(n)}\right)\right)^{m_1}. \end{aligned}$$

Thus by choosing $m_1 = O(d^2 c^4 T^2(n) \log(n/d))$, we can ensure that $p_f = o(1)$, and hence, M is d-disjunct with overwhelming probability. The agument for (d, e_1)-disjunctness is similar, but uses Chernoff bounds to upper bound the failure probability that $X \leq e_1$ for some choice of v and A. ∎

Remark 9: In Design 1, we need to assume that the designated vertices (if any) are not defective, and hence, their corresponding columns can be removed from the matrix M. By doing so, we will be able to ensure that the resulting matrix is disjunct. Obviously, such a restriction cannot be avoided since, for example, M might be forced to contain an all-ones column corresponding to one of the designated vertices and thus, fail to be even 1-disjunct.

Remark 10: By applying Theorem 8 on the complete graph (using Design 1), we get $O(d^2 \log(n/d))$ measurements, since in this case, the mixing time is $T(n) = 1$. Thereby, we recover the trade-off obtained by the probabilistic construction in classical group testing (note that classical group testing corresponds to graph-constrained group testing on the vertices of the complete graph).

Remark 11: By considering the special case of complete graphs it is possible to establish that the cubic dependence of m_3 on d and dependence of m_4 on the degree parameter D are necessary and cannot be in general improved. We omit the details due to space constraints.

Now we instantiate the result obtained in Theorem 8 to two important special cases; namely, expander graphs and Erdős-Rényi random graph $G(n,p)$. We will use the following results (proofs can be found in the full version of the paper):

Lemma 12: Suppose that G is a regular expander graph with constant spectral gap. Then $T(n) = O(\log n)$. ∎

Lemma 13: For every $\epsilon > 0$, there is a constant $\alpha > 0$ such that the following holds. Suppose that G is a random graph $G(n,p)$ with average degree $np \geq \alpha \ln n$. Then, with probability $1 - o(1)$,

- The graph G is $(np(1-\epsilon), (1+\epsilon)/(1-\epsilon))$-uniform,
- We have $T(n) = O(\log n)$. ∎

Using the above lemmas in Theorem 8, we get the following result:

Theorem 14: There is an integer $D_0 = \Omega(d \log^2 n)$ such that for every $D \geq D_0$ the following holds: Suppose that the graph G is either

1) A D-regular expander graph with constant spectral gap, or,
2) A random graph $G(n, D/n)$.

Then for every $p \in [0,1)$, with probability $1 - o(1)$ Designs 1, 2, 3, and 4 output (d,e)-disjunct matrices (not considering the columns corresponding to the designated vertices and the sink in Designs 1 and 3), for some $e = \Omega(dp \log n)$, using respectively m_1, m_2, m_3, m_4 measurements, where $m_1, m_2 = O(d^2 \log^3 n)$, $m_3 = O(d^3 \log^5 n)$, and $m_4 = O(d^3 D \log^5 n)$. ∎

Decoding: In light of Theorems 8 and 14, we know that the matrix M output by our proposed designs is almost surely guaranteed to be (k,e)-disjunct, and moreover, is consistent with the underlying graph G. Therefore, as discussed in Section II, the set of pools defined by the rows of M can be used to distinguish any set of up to d defective vertices (or edges) in G. We further point out that, as in the case of classical group testing, the set of defective items (i.e., defective vertices or edges of G) can be reconstructed using the following (well known) simple and efficient decoding procedure: Let $y \in \{0,1\}^m$ denote the vector consistintg of the measurement outcomes. Then an item i is labeled as defective iff the ith column of M, denoted by C_i, satisfies $|C_i \setminus y| \leq e$. See [11] for a detailed discussion on this method.

REFERENCES

[1] R. Dorfman, "The detection of defective members of large populations," *Annals of Mathematical Statistics*, vol. 14, pp. 436–440, 1943.

[2] D.-Z. Du and F. Hwang, *Combinatorial Group Testing and its Applications*, 2nd ed. World Scientific Publishing Company, 2000.

[3] G. Atia and V. Saligrama, "Noisy group testing: An information theoretic perspective," in *Proceedings of Allerton, UIUC*, 2009.

[4] M. Sobel and P. Groll, "Group testing to eliminate efficiently all defectives in a binomial sample," *Bell Systems Technical Journal*, vol. 38, pp. 1179–1252, 1959.

[5] P. Pevzner and R. Lipshutz, "Towards DNA sequencing chips," in *Proc. of MFCS*, ser. LNCS, vol. 841, 1994, pp. 143–158.

[6] A. Blass and Y. Gurevich, "Pairwise testing," *Bulletin of the EATCS*, vol. 78, pp. 100–132, 2002.

[7] J. Wolf, "Born-again group testing: multiaccess communications," *IEEE Transactions on Information Theory*, vol. 31, pp. 185–191, 1985.

[8] N. Duffield, "Network tomography of binary network performance characteristics," *IEEE Trans. on Inf. Theory*, vol. 52, no. 12, pp. 5373–5388, 2006.

[9] H. Nguyen and P. Thiran, "The boolean solution to the congested IP link location problem: Theory and practice," in *Proc. of INFOCOM)*, 2007, pp. 2117–2125.

[10] ——, "Using end-to-end data to infer lossy links in sensor networks," in *Proc. of INFOCOM*, 2006.

[11] M. Cheraghchi, A. Hormati, A. Karbasi, and M. Vetterli, "Compressed sensing with probabilistic measurements: A group testing solution," in *Proceedings of Allerton, UIUC*, 2009.

[12] M. Aigner, *Combinatorial Search*. New York: Wiley-Teubner Series in Computer Science, Wiley, 1988.

[13] D. B. Wilson, "Generating random spanning trees more quickly than the cover time," in *Proc. of STOC*, 1996, pp. 296–303.