Non-adaptive Group Testing: Explicit bounds and novel algorithms

Chun Lam Chan*, Sidharth Jaggi*, Venkatesh Saligrama+, and Samar Agnihotri*

The Chinese University of Hong Kong*, Boston University+

***Abstract*—THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD[1]. We present computationally efficient and provably correct algorithms with near-optimal sample-complexity for noisy non-adaptive group testing. Group testing involves grouping arbitrary subsets of items into pools. Each pool is then tested to identify the defective items, which are usually assumed to be sparsely distributed. We consider random non-adaptive pooling where pools are selected randomly and independently of the test outcomes. Our noisy scenario accounts for both false negatives and false positives for the test outcomes. Inspired by compressive sensing algorithms we introduce four novel computationally efficient decoding algorithms for group testing, CBP via Linear Programming (CBP-LP), NCBP-LP (Noisy CBP-LP), and the two related algorithms NCBP-SLP+ and NCBP-SLP- ("Simple" NCBP-LP). The first of these algorithms deals with the noiseless measurement scenario, and the next three with the noisy measurement scenario. We derive explicit sample-complexity bounds—with all constants made explicit—for these algorithms as a function of the desired error probability; the noise parameters; the number of items; and the size of the defective set (or an upper bound on it). We show that the sample-complexities of our algorithms are near-optimal with respect to known information-theoretic bounds.**

I. INTRODUCTION

The goal of *group testing* is to identify a small unknown subset $\mathcal{D}$ of defective items embedded in a much larger set $\mathcal{N}$ (usually in the setting where $d = |\mathcal{D}|$ is much smaller than $n = |\mathcal{N}|$, *i.e.*, d is $o(n)$). This problem was first considered by Dorfman [1] in scenarios where multiple items in a group can be simultaneously tested, with a binary output depending on whether or not a "defective" item is present in the group being tested. In general, the goal of group testing algorithms is to identify the defective set with as few measurements as possible. As demonstrated in [1] and later work [2], with judicious grouping and testing, far fewer than the trivial upper bound of n tests may be required to identify $\mathcal{D}$.

We consider *non-adaptive group testing* in this paper. In non-adaptive group testing, the set of items being tested in each test is required to be independent of the outcome of every other test [2]. This restriction is often useful in practice, since this enables parallelization of the testing process. It also allows for an automated testing process. In contrast, the procedures and hardware required for *adaptive* group testing may be significantly more complex.

In this paper we describe computationally efficient algorithms with near-optimal performance for noiseless and noisy non-adaptive group testing problems. We describe the different aspects of the paper in some detail next.

"Noisy" measurements: In addition to the *noiseless* group-testing problem, we consider the "noisy" variant of the problem. In this noisy variant the result of each test may differ from the true result (in an independent and identically distributed manner) with a certain pre-specified probability q. This leads to both false positives and negatives in the test outcomes. Much of the existing work either considers one-sided noise, namely false positives [3] but no false negatives or a "worst-case" noise [4] wherein the number of false positives and negatives are assumed bounded.[2] Since the measurements are noisy, the problem of estimating the set of defective items is more challenging, and is known to require more tests.[3] The work closest to this work is [7], where explicit upper and lower bounds for the group-testing problem were first derived.

Computationally efficient and near-optimal algorithms: Most algorithms in the literature focus on optimizing the number of measurements required – in some cases, this leads to algorithms that may not be computationally efficient to implement (for *e.g.* [3]). In this paper we present algorithms that are not only computationally efficient but are also near-optimal in the number of measurements required.

We analyze three different types of algorithms (the last one has two flavors), related to those described in the compressive sensing literature (see Section I-A).

Our algorithms are related to linear programming relaxations used in the compressive sensing literature. In compressive sensing the ℓ_0 norm minimization is relaxed to an ℓ_1 norm minimization. In the noise-free case this relaxation results in a linear program since the measurements are linear. In contrast, in group testing, the measurements are non-linear

[1]The first author is an undergraduate student, and this material arose from his senior thesis.

[2]For instance [4] considers group-testing algorithms that are resilient to *all* noise patterns wherein at most a fraction q of the results differ from their true values, rather than the probabilistic guarantee we give against *most* fraction-q errors. This is analogous to the difference between combinatorial coding-theoretic error-correcting codes (for instance Gilbert-Varshamov codes [5]) and probabilistic information-theoretic codes (for instance [6]). In this work we concern ourselves only with the latter, though it is possible that our techniques can also be used to analyzed the former.

[3]**We wish to highlight the difference between *noise* and *errors*. We use the former term to refer to noise in the outcomes of the group-test, regardless of the group-testing algorithm used. The latter term is used to refer to the error due to the estimation process of the group-testing algorithm.**

and boolean. In the noise-free case the measurements take the value one if *some* defective item is in the pool and zero if no defective item is part of the pool. Furthermore, noise in the group testing scenario is also boolean unlike additive noises in compressive sensing. For these reasons we also need to relax our boolean measurement equations. We do so by using a novel combination of inequality and positivity constraints. Our LP formulation and analysis is related to error-correction [8], where, one uses a "minimum distance" decoding criteria based on perturbation analysis. The idea is to decode to a vector pair consisting of defective items, $\mathbf{x}$, and the error vector, η such that the error-vector η is as "small" as possible by solving a sequence of LPs. We call this algorithm the Noisy Combinatorial Basis Pursuit via LP decoding (NCBP-LP). Using standard concentration results we show that the solution to our LP decoding algorithm recovers the true defective items with high probability. Furthermore, we achieve near-optimal performance in the sense that our sample complexity for NCBP-LP match the lower bounds within a constant factor, where the constant is independent of the number of items n and the defective set size d (but may depend on the noise parameter q, and the error probability ϵ). Based on this analysis, we can directly derive the performance of two other LP-based decoding algorithms. In particular CBP-LP considers the noiseless measurement scenario, and NCBP-SLP+ and NCBP-SLP- consider the noisy measurement scenario, but *only* use constraints corresponding to positive and negative test outcomes respectively.

"Small-error" probability: Existing work has considered both deterministic and random pooling designs [2]. In this context both deterministic and probabilistic sample complexity bounds for the number of measurements T that lead to exact identification of the defective items have been derived. There is also existing work on characterizing sample-complexity bounds for the average case scenario (see [3]). These sample complexity bounds are usually asymptotic in nature and describe the scaling of the number of items n with respect to the number of defectives d to ensure that the error probability approaches zero. To gain new insights into the constants involved in the sample-complexity bounds we admit a small but fixed error probability, ϵ. With this new perspective we can derive upper bounds that hold not only in an order-wise sense but also where the constants involved in these order-wise bounds can be made explicit.

Explicit Sample Complexity Bounds: Our sample complexity bounds are of the form $T \geq \beta(q,\epsilon)d\log(n)$. The function $\beta(q,\epsilon)$ is an explicitly computed function of the noise parameter q and admissible error probability ϵ. In the literature, order-optimal upper and lower bounds on the number of tests required are known for the problems we consider (for instance [3], [9]). In both the noiseless and noisy variants, the number of measurements required to identify the set of defective items is known to be $T = \Theta(d\log(n))$ – here $n = |\mathcal{N}|$ is the total number of items and $d = |\mathcal{D}|$ is the size of the defective subset. In fact, if only D, an upper bound on d, is known, then $T = \Theta(D\log(n))$ measurements are also known to be necessary and sufficient. In our algorithms we explicitly demonstrate that we require only a knowledge of D rather than the exact value of d. Furthermore, in the noisy variant, we show that the number of tests required is in general at most a constant factor larger than in the noiseless case (where this constant β is independent of both n and d, but may depend on the noise parameter q and the allowable *error-probability* ϵ of the algorithm).

This paper is organized as follows. In Section II, we introduce the model and corresponding notation, and describe the algorithms analyzed in this work. In Section III, we describe the main results of this work. Section IV contains the analysis of the group-testing algorithms considered.

A. Compressive Sensing

Compressive sensing has received significant attention over the last decade. We describe the version most related to the topic of this paper [10], [11]. This version considers the following problem. Let $\mathbf{x}$ be an *exactly d-sparse* vector in $\mathbb{R}^n$, *i.e.*, a vector with at most d non-zero components (in general in the situations of interest $d = o(n)$).

Let $\mathbf{z}$ correspond to a *noise vector* added to the measurement $M\mathbf{x}$. One is given a set of "compressed noisy measurements" of $\mathbf{x}$ as $\mathbf{y} = M\mathbf{x} + \mathbf{z}$. (The noise is guaranteed to be not "too large" ($||\mathbf{z}||_2 \leq c_2$).) The $T \times n$ matrix M is designed by choosing each entry i.i.d. from a suitable probability distribution (for instance, the set of zero-mean, $1/n$ variance Gaussian random variables). The decoder must use the resulting *noisy measurement vector* $\mathbf{y} \in \mathbb{R}^T$ and the matrix M to estimate in a computationally efficient manner the underlying vector $\mathbf{x}$. The challenge is to do so with as few measurements as possible, *i.e.*, with the number of rows T of M being as small as possible.

1) Basis Pursuit: An alternate decoding procedure proceeds by relaxing the compressive sensing problem into the convex optimization problem called *Basis Pursuit* (BP).

$$\mathbf{x} = \arg\min ||\mathbf{x}||_1 \quad (1)$$

$$\text{subject to } ||\mathbf{y} - M\mathbf{x}||_2 \leq c_2 \quad (2)$$

It can be shown (for instance [10], [11]) that there exist constants c_4, c_5 and c_6 such that with $T = c_4 d\log(n)$, with probability at least $1 - 2^{c_5 n}$, the solution $\mathbf{x}^*$ to BP satisfies $||\mathbf{x}^* - \mathbf{x}||_2 \leq c_6||\mathbf{z}||_2$.

II. Background

A. Model and Notation

A set $\mathcal{N}$ contains n items, of which an unknown subset $\mathcal{D}$ are said to be "defective".[4] The goal of group-testing is to correctly identify the set of defective items via a minimal number of "group tests", as defined below.

[4]It is common (see for example [12]) to assume that the number d of defective items in $\mathcal{D}$ is known, or at least a good upper bound D on d, is known *a priori*. If not, other work (for example [13]) considers non-adaptive algorithms with low query complexity that help estimate d. However, in this work we explicitly consider algorithms that do not require such foreknowledge of d – rather, our algorithms have "good" performance with $\mathcal{O}(D\log(n))$ measurements.

Each row of a $T \times n$ binary *group-testing matrix* M corresponds to a distinct test, and each column corresponds to a distinct item. Hence the items that comprise the group being tested in the ith test are exactly those corresponding to columns containing a 1 in the ith location. The method of generating such a matrix M is part of the design of the group test. This along with the method of estimating the set $\mathcal{D}$, is described in Section II-B.

The length-n binary *input* vector $\mathbf{x}$ represents the set $\mathcal{N}$, and contains 1s exactly in the locations corresponding to the items of $\mathcal{D}$. The locations with ones/defective items are said to be *positive* – the other locations are said to be *negative*. We use these terms interchangeably.

The outcomes of the *noiseless* tests correspond to the length-T binary *noiseless result* vector $\mathbf{y}$, with a 1 in the i^{th} location if and only if the ith test contains at least one defective item.

The observed vector of test outcomes in the *noisy* scenario is denoted by the length-T binary *noisy result* vector $\hat{\mathbf{y}}$ – the probability that each entry y_i of $\mathbf{y}$ differs from the corresponding entry $\hat{y}_i$ in $\hat{\mathbf{y}}$ is q, where q is the *noise parameter*. The locations where the noiseless and the noisy result vectors differ is denoted by the length-T binary *noise vector* ν, with 1s in the locations where they differ.

The estimate of the locations of the defective items is encoded in the length-n binary *estimate vector* $\hat{\mathbf{x}}$, with 1s in the locations where the group-testing algorithms described in Section II-B estimate the defective items to be.

The *probability of error* of any group-testing algorithm is defined as the probability (over the input vector $\mathbf{x}$, group-testing matrix M, and noise vector ν) that the estimated vector differs from the input vector.

B. Algorithms

We now describe our algorithms in both the noiseless and noisy settings. The algorithms are specified by the choices of encoding matrices and decoding algorithms. The $T \times n$ group-testing matrix M is defined by randomly selecting each entry in it in an i.i.d. manner to equal 1 with probability $p = 1/D$, and 0 otherwise.

Noisy CBP-LP (NCBP-LP):

A linear relaxation of the group-testing problem leads naturally to NCBP-LP (3-9). In particular, each x_i is relaxed to satisfy $0 \le x_i \le 1$, $\bar{d}$ represents our "guess" for the value of $d \le D$, and the non-linear measurements are linearized in (4)-(5). Also, we define "slack" variables η_i for all $i \in \{1, \ldots, T\}$ to account for errors in the test outcome. For a particular test i this η_i is defined to be zero if a particular test result is correct, and positive (and at least 1) if the test result is incorrect. Of course, the decoder does not know *a priori* which scenario a particular test outcome falls under, and hence has to also decode η. Nonetheless, as is common in the field of error-correction [8], often using a "minimum distance" decoding criteria (decoding to a vector pair $(\mathbf{x}, \eta)$ such that the error-vector η is as "small" as possible) leads to good decoding performance. Our LP decoder attempts to do so.

$$\forall \bar{d} \in \{0, \ldots, D\},\ (\hat{\mathbf{x}}(\bar{d}), \hat{\eta}(\bar{d})) = \arg\min_{\mathbf{x}, \eta} \sum_i \eta_i \tag{3}$$

such that

$$-\eta_i + \sum_{j: m_{ij}=1} x_j = 0, \text{ if } \hat{y}_i = 0, \tag{4}$$

$$\eta_i + \sum_{j: m_{ij}=1} x_j \ge 1, \text{ if } \hat{y}_i = 1, \tag{5}$$

$$\sum_{\forall j} x_j = \bar{d} \le D \tag{6}$$

$$0 \le x_j \le 1. \tag{7}$$

$$0 \le \eta_i \le D, \text{ if } \hat{y}_i = 0, \tag{8}$$

$$0 \le \eta_i \le 1, \text{ if } \hat{y}_i = 1, \tag{9}$$

Fig. 1. **NCBP-LP:** Constraint (7) relaxes the constraint that each $x_j \in \{0, 1\}$, and constraint (6) indicates that there are exactly $\bar{d}$ defective items in the $\bar{d}$th iteration of the LP. The variables η_i are "slack variables" in the equations (4) and (5). For instance, if test i is truly negative, then all the variables in an equation of the form (4) are zero. However, if the test is a false negative, then the variable η_i is then set to equal the number of defective items in test i. Similarly, if a test i is truly positive, then η_i is zero, and equations of the form (5) are satisfied. However, if the test is a false positive, then η_i is set to equal 1 (and the x_j variables tested in test j are set to 0). Note that η_i is bounded above by 1 in the case of (false) positives, but is only bounded above by D in the case of (false) negatives. This is due to the asymmetry of positive and negative test outcomes – multiple positive items tested simultaneously do not give a different outcome from a single positive item tested.

$$\forall \bar{d} \in \{0, \ldots, D\},\ \hat{\mathbf{x}}(\bar{d}) = \text{ feasible point in}$$

$$\sum_{j: m_{ij}=1} x_j = 0, \text{ if } y_i = 0, \tag{10}$$

$$\sum_{j: m_{ij}=1} x_j \ge 1, \text{ if } y_i = 1, \tag{11}$$

$$\sum_{\forall i} x_j = \bar{d} \le D \tag{12}$$

$$0 \le x_j \le 1 \tag{13}$$

Fig. 2. **CBP-LP:** This LP simply attempts to find *any* feasible solution for any value of $\bar{d} \in \{0, \ldots, D\}$

To be more precise, we define the pair $(\hat{\mathbf{x}}, \hat{\eta})$ to be *feasible* if $\hat{\mathbf{x}}$ is a binary vector of weight at most D, and $\hat{\eta} \in \mathbb{R}^T$ is a vector of Hamming weight at most $Tq(1+\tau)$. (The value of τ is a code-design parameter to be specified later.) We then solve the sequence of LPs in (3–9) for each $\bar{d} \in \{1, \ldots, D\}$ and output any feasible pair $(\hat{\mathbf{x}}(\bar{d}), \hat{\eta}(\bar{d}))$ in the sequence (and output an error if there are none, or more than one).

The decoder sequentially attempts to find valid solutions for the above LP for sequentially increasing integers starting from $\bar{d} = 0$. If no feasible solution is found, or if an infeasible solution is found, the decoder increments $\bar{d}$ by 1 and continues until it reaches $\bar{d} = D$. If a valid solution is found for any value of $\bar{d}$, the decoder stops and outputs that. If no valid solution is found for any value of $\bar{d} \le D$, the decoder declares an error. Our analysis demonstrates that this algorithm has a "small" probability of error.

Combinatorial Basis Pursuit via LP decoding (CBP-LP):

CBP-LP, which analyzes the scenario with noiseless mea-

$$\forall \bar{d} \in \{0, \ldots, D\},, \qquad \forall \bar{d} \in \{0, \ldots, D\},$$

$$(\hat{\mathbf{x}}(\bar{d}), \hat{\eta}(\bar{d})) = \arg\min_{\mathbf{x},\eta} \sum_{i:\hat{y}_i=1} \eta_i \qquad (\hat{\mathbf{x}}(\bar{d}), \hat{\eta}(\bar{d})) = \arg\min_{\mathbf{x},\eta} \sum_{i:\hat{y}_i=0} \eta_i \tag{14}$$

such that $\qquad$ such that

$$\eta_i + \sum_{j:m_{ij}=1} x_j \geq 1, \text{ if } \hat{y_i} = 1, \qquad -\eta_i + \sum_{j:m_{ij}=1} x_j = 0, \text{ if } \hat{y_i} = 0, \tag{15}$$

$$\sum_{\forall j} x_j = \bar{d} \leq D, \qquad \sum_{\forall j} x_j = \bar{d} \leq D \tag{16}$$

$$0 \leq x_j \leq 1, \qquad 0 \leq x_j \leq 1 \tag{17}$$

$$0 \leq \eta_i \leq 1, \text{ if } \hat{y_i} = 1, \qquad 0 \leq \eta_i \leq D, \text{ if } \hat{y_i} = 0, \tag{18}$$

NCBP-SLP+ $\qquad$ **NCBP-SLP-**

Fig. 3. **NCBP-SLP+, NCBP-SLP-:** The variables in the objective functions and the set of constraints for both these LPs are subsets of those in NCBP-LP.

surements, is a special case of NCBP-LP with each η_j variable set to zero. Hence it reduces to the problem of finding *any* feasible point in the constraint set (10-13)

NCBP via Simpler LP decoding (NCBP-SLP):

In fact, it turns out that simpler LPs still gives essentially the same performance as NCBP-LP. Consider the two LPs given above. The intuition is that if NCBP-LP and works by finding a η vector with low Hamming weight, then NCBP-SLP+ (respectively NCBP-SLP-) does the same by finding a η vector with low Hamming weight restricted just to the set of positive (respectively negative) outcomes. Since the noise that converts $\mathbf{y}$ to $\hat{\mathbf{y}}$ is probabilistic, by standard concentration results these two approaches should, with high probability, lead to the same result.

III. Main Results

The analysis of the constants in the three main theorems are not optimized[5] but the constants are given to demonstrate the functional dependence on δ and q. Our algorithms' sample complexities are commensurate (up to a constant factor) with information-theoretic lower bounds of $\Omega(D\log(n))$ tests [7].

We define Γ as $\ln(d)/\ln(n)$ and γ as $(\Gamma+\delta)/(1+\delta)$ (note that in the limit of large n, Γ lies in the interval $[0, 1)$ and γ in the interval $(\delta/(\delta+1), 1]$.

Theorem 1: NCBP-LP with error probability at most $n^{-\delta}$ requires no more than $\beta_{LP} D \log n$ tests, with β_{LP} defined as

$$\max\left\{\frac{4e\,(\delta+1+\Gamma)}{(1-2q)^2}, 8e\,(\delta+1+\Gamma), \frac{4(1-q+2qe)e\,(\delta+1+\Gamma)}{(1-q)^2}, \frac{8e\,(\delta+1+\Gamma)}{(1-q+2qe)}, \frac{16(1+\sqrt{\gamma})^2(1+\delta)\ln 2}{(1-e^{-2})(1-2q)^2}\right\}.$$

Theorem 2: CBP-LP with error probability at most $n^{-\delta}$ requires no more than $\frac{16(1+\sqrt{\gamma})^2(1+\delta)\ln 2}{(1-e^{-2})(1-2q)^2} D \log n$ tests.

Theorem 3: NCBP-SLP+ and NCBP-SLP- with error probability at most $n^{-\delta}$ require no more than $\beta_{SLP+} D \log n$ and $\beta_{SLP-} D \log n$ tests respectively, with

$$\beta_{SLP+} = \max\left\{\frac{4e\,(\delta+1+\Gamma)}{(1-2q)^2}, 8e\,(\delta+1+\Gamma), \frac{16(1+\sqrt{\gamma})^2(1+\delta)\ln 2}{(1-e^{-2})(1-2q)^2}\right\},$$

$$\beta_{SLP-} = \max\left\{\frac{4(1-q+2qe)e\,(\delta+1+\Gamma)}{(1-q)^2}, \frac{8e\,(\delta+1+\Gamma)}{(1-q+2qe)}, \frac{16(1+\sqrt{\gamma})^2(1+\delta)\ln 2}{(1-e^{-2})(1-2q)^2}\right\}$$

[5] Doing so is analytically very cumbersome.

IV. Analysis of Algorithms

We first sketch the proof of Theorem 1 (full details are in [14]), and note that Theorems 2 and 3 are direct corollaries.

Proof sketch of Theorem 1: Without loss of generality, let $\mathbf{x}$ be the vector with 1s in the first d locations, and 0s in the last $n-d$ locations.[6]

At a high level, our proof proceeds as follows. First, we invoke Theorem 4 from [7] that demonstrates that for appropriate choice of parameters τ and T in fact, with high probability, there is exactly one value of $\bar{d} \leq D$ for which a feasible pair $(\hat{\mathbf{x}}, \hat{\eta}(\bar{d}))$ exists (and this pair is unique), and in fact this $\bar{d} = d$. The proof in [14] is information-theoretic in spirit, and is analogous to "nearest-neighbour decoding". Hence in our proof we can focus on simply this value of d.

Next, we define a *finite* set $\mathbf{\Phi'}$ (that depends on the true $\mathbf{x}$) containing so-called "perturbation vectors".[7] In particular, $\mathbf{\Phi'} = \{\phi'\}_{k'=1}^{d(n-d)}$ is the set of $d(n-d)$ vectors with a single -1 in the support of $\mathbf{x}$, a single 1 outside the support of $\mathbf{x}$, and zeroes everywhere else. For instance, the first ϕ' in the set equals $(-1, 0, \ldots, 0, 1, 0, \ldots, 0)$, where the 1 is in the $(d+1)$th location. We demonstrate that any $\bar{\mathbf{x}}$ in the feasible set of the constraint set of NCBP-LP can be written as the true $\mathbf{x}$ plus a *non-negative linear combination* of perturbation vectors from this set. The physical intuition behind the proof is that the vectors from $\mathbf{\Phi'}$ correspond to a "mass-conserving" perturbation of $\mathbf{x}$. The property of non-negativity of the linear combinations arises from a physical argument demonstrating that there is a path from $\mathbf{x}$ to any point in the feasible set using these perturbations, over which one never has to "back-track". The non-negative linear combination property is important, since this enables us to characterize the directions in which a vector can be perturbed from $\mathbf{x}$ to another vector that satisfies the constraints of NCBP-LP, in a "finite" manner (instead of having to consider the uncountably infinite number of directions that $\mathbf{x}$ could be perturbed to). The non-negativity of the linear combination is also crucial since, as we explain below, this property ensures that the objective function of the LP can only increase when perturbed in a convex combination of the directions in $\mathbf{\Phi'}$.

The heart of our argument then lies in the characterization with an exhaustive[8] case-analysis of the expected change (over randomness in the matrix M and noise ν) in the value of each slack variable η_i when $\mathbf{x}$ is perturbed to some $\bar{\mathbf{x}}$ by a vector in $\mathbf{\Phi'}$. In particular, we demonstrate that for each such *individual* perturbation vector, the expected change in the value of each slack variable η_i is *strictly* positive with high probability. The actual proof follows from a case-analysis similar to the one performed in the example in Table I.

With slightly careful use of standard concentration inequalities (specifically, we need to use both the additive and multiplicative forms of the Chernoff bound) we show that the

[6] As can be verified, our analysis is agnostic to the actual choice of $\mathbf{x}$, as long as it is a vector in $\{0, 1\}^n$ of weight any $d \leq D$.

[7] This set is defined just for the purpose of this proof – the encoder/decoder do not need to know this set.

[8] And exhausting!

						$\mathbf{x}$ $(\mathbf{1},\mathbf{1},\mathbf{0})$	$\mathbf{x}' = \mathbf{x} + \phi'$ $(\mathbf{0},\mathbf{1},\mathbf{1})$	
1. $\hat{\mathbf{y}}_i$	2. $\eta(\mathbf{x})$	3. $\mathbf{y}_i$	4. $\mathbf{m}_i$	5. $P(\hat{\mathbf{y}}_i, \mathbf{m}_i \vert \mathbf{x})$	6. $\eta_i(\mathbf{m}_i, \mathbf{x})$	7. $\eta_i(\mathbf{m}_i, \mathbf{x})$	8(a). $\eta_i(\mathbf{m}_i, \mathbf{x}')$	8(b). $\mathrm{E}(\mathbf{m}_i, \Delta'_i)$
1	$(1 - \mathbf{m}_i.\mathbf{x})^+$	0	$(\mathbf{0},\mathbf{0},\mathbf{0})$	$q(1-p)^3$	1	1	1	0
			$(\mathbf{0},\mathbf{0},\mathbf{1})$	$q(1-p)^2p$	$(1-x_3)^+$	1	0	$-q(1-p)^2p$
		1	$(\mathbf{0},\mathbf{1},\mathbf{0})$	$(1-q)(1-p)^2p$	$(1-x_2)^+$	0	0	0
			$(\mathbf{0},\mathbf{1},\mathbf{1})$	$(1-q)(1-p)p^2$	$(1-x_2-x_3)^+$	0	0	0
			$(\mathbf{1},\mathbf{0},\mathbf{0})$	$(1-q)(1-p)^2p$	$(1-x_1)^+$	0	1	$(1-q)(1-p)^2p$
			$(\mathbf{1},\mathbf{0},\mathbf{1})$	$(1-q)(1-p)p^2$	$(1-x_1-x_3)^+$	0	0	0
			$(\mathbf{1},\mathbf{1},\mathbf{0})$	$(1-q)(1-p)p^2$	$(1-x_1-x_2)^+$	0	0	0
			$(\mathbf{1},\mathbf{1},\mathbf{1})$	$(1-q)p^3$	$(1-x_1-x_2-x_3)^+$	0	0	0
								$(1-2q)(1-p)^2p$
0	$\mathbf{m}_i.\mathbf{x}$	0	$(\mathbf{0},\mathbf{0},\mathbf{0})$	$(1-q)(1-p)^3$	0	0	0	0
			$(\mathbf{0},\mathbf{0},\mathbf{1})$	$(1-q)(1-p)^2p$	x_3	0	1	$(1-q)(1-p)^2p$
		1	$(\mathbf{0},\mathbf{1},\mathbf{0})$	$q(1-p)^2p$	x_2	1	1	0
			$(\mathbf{0},\mathbf{1},\mathbf{1})$	$q(1-p)p^2$	x_2+x_3	1	2	$q(1-p)p^2$
			$(\mathbf{1},\mathbf{0},\mathbf{0})$	$q(1-p)^2p$	x_1	1		
			$(\mathbf{1},\mathbf{0},\mathbf{1})$	$q(1-p)p^2$	x_1+x_3	1	0	0
			$(\mathbf{1},\mathbf{1},\mathbf{0})$	$q(1-p)p^2$	x_1+x_2	2	1	$-q(1-p)p^2$
			$(\mathbf{1},\mathbf{1},\mathbf{1})$	qp^3	$x_1+x_2+x_3$	2	2	0
								$(1-2q)(1-p)^2p$

TABLE I. Suppose $\mathbf{x} = (1,1,0)$. Choose some $\mathbf{x}' \neq \mathbf{x}$ (in this example, $\mathbf{x}' = \mathbf{x} + \phi'$, where $\phi' = (-1,0,1)$ is a *perturbation vector*). This example analyzes the expectation (over the randomness in the particular row $\mathbf{m}_i$ of the measurement matrix M) of the difference in value of the corresponding slack variables $\eta_i(\mathbf{x})$ and $\eta_i(\mathbf{x}')$ in column $8(b)$. To compute these, we consider the columns of the table above sequentially from left to right. Column 1 considers the two possible values of the observed vector $\hat{y}_i$. Column 2 gives the corresponding values of the slack variables corresponding to the ith test, as returned by the constraints (4) and (5) of NCBP-LP – here $(f(\mathbf{x}))^+$ denotes the function $\max\{f(\mathbf{x}), 0\}$. Column 3 indexes the possibilities of the (noiseless) test outcomes y_i, and column 4 enumerates possible values for $\mathbf{m}_i$, the i-th row of M, that could have generated the values of y_i in column 3, given that $\mathbf{x} = (1,1,0)$. Column 5 computes the probability of a particular observation $\hat{y}_i$ and a row $\mathbf{m}_i$, given that the noiseless output y_i equaled a particular value. Column 6 computes the function in column 2, given that $\mathbf{m}_i$ equals the value given in Column 4. Columns 7 and $8(a)$ respectively explicitly compute the value of the function in column 6 for the vectors $\mathbf{x}$ and $\mathbf{x}'$ – the red entries in column $8(a)$ index those locations where $\eta(\mathbf{x}')$, the slack variable for the perturbed vector, is less than $\eta(\mathbf{x})$, and the green cells indicate those locations where the situation is reverse. Column $8(b)$ then computes the product of column 5 with the difference of the entries in column 7 from those of column $8(a)$, *i.e.*, the expected change in the value of the slack variable $\eta_i(.)$. The value $(1-2q)(1-p)^2p$ in blue at the bottom represents the expected change (averaged over all possible tuples $(y_i, \mathbf{m}_i, \hat{y}_i)$).

probability distributions derived above concentrate. We then take the union bound over all vectors in $\mathbf{\Phi}'$ (in fact, there are a total of $d(n-d)+d$ such vectors in $\mathbf{\Phi}'$) and show that with high probability the expected change in the value of the objective function (which equals the weighted sum of the changes in the values of the slack variables η_i) for *each* perturbation vector in $\mathbf{\Phi}'$ is also strictly positive.

Finally, we note that the set of feasible $(\bar{\mathbf{x}}, \eta)$ of NCBP-LP forms a convex set. Hence if η strictly increases along every direction in $\mathbf{\Phi}'$, then in fact η strictly increases when the true $\mathbf{x}$ is perturbed in *any* direction (since, as noted before, any vector in the feasible set can be written as $\mathbf{x}$ plus a non-negative linear combination of vectors in Φ'). Hence the true $\mathbf{x}$ must be the solution to NCBP-LP. ■

Proof of Theorem 2: We substitute $q = 0$ into Theorem 1 and choose the largest term. ■

Proof of Theorem 3: The proof is essentially the same as in the case of Theorem 1. Details in [14]. ■

References

[1] R. Dorfman, "The detection of defective members of large populations," *Annals of Mathematical Statistics*, vol. 14, pp. 436–411, 1943.

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

[3] G. Atia and V. Saligrama, "Boolean compressed sensing and noisy group testing," *CoRR*, vol. abs/0907.1061, 2009.

[4] A. J. Macula, "Error-correcting nonadaptive group testing with de-disjunct matrices," *Discrete Applied Mathematics*, vol. 80, no. 2-3, pp. 217 – 222, 1997.

[5] E. N. Gilbert, "A comparison of signaling alphabets," *Bell System Technical Journal*, vol. 31, pp. 504–522, 1952.

[6] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, pp. 379–423, Jul. 1948.

[7] C. L. Chan, P. H. Che, S. Jaggi, and V. Saligrama, "Non-adaptive probabilistic group testing with noisy measurements: Near-optimal bounds with efficient algorithms," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, Oct. 2010.

[8] F. J. McWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. North-Holland, 1977.

[9] D. Sejdinovic and O. Johnson, "Note on noisy group testing: Asymptotic bounds and belief propagation reconstruction," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, Oct. 2010, pp. 998 –1003.

[10] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[11] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, December 2008.

[12] A. Macula, "Probabilistic nonadaptive group testing in the presence of errors and dna library screening," *Annals of Combinatorics*, vol. 3, pp. 61–69, 1999.

[13] M. Sobel and R. M. Elashoff, "Group testing with a new goal, estimation," *Biometrika*, vol. 62, no. 1, pp. 181–193, 1975.

[14] 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.