A MaxSAT-based Framework for Group Testing

Lorenzo Ciampiconi,[1] **Bishwamittra Ghosh,**[2] **Jonathan Scarlett,**[2] **Kuldeep S Meel**[2]

[1]Politecnico di Milano, [2]National University of Singapore

Abstract

The success of MaxSAT (maximum satisfiability) solving in recent years has motivated researchers to apply MaxSAT solvers in diverse discrete combinatorial optimization problems. Group testing has been studied as a combinatorial optimization problem, where the goal is to find defective items among a set of items by performing sets of tests on items. In this paper, we propose a MaxSAT-based framework, called MGT, that solves group testing, in particular, the decoding phase of non-adaptive group testing. We extend this approach to the noisy variant of group testing, and propose a compact MaxSAT-based encoding that guarantees an optimal solution. Our extensive experimental results show that MGT can solve group testing instances of 10000 items with 3% defectivity, which no prior work can handle to the best of our knowledge. Furthermore, MGT has better accuracy than the LP-based approach. We also discover an interesting phase transition behavior in the runtime, which reveals the easy-hard-easy nature of group testing.

1 Introduction

Given a large set of items containing a subset of defective items, the problem of group testing concerns the design and evaluation of tests on pools of items (a pool is a selected subset of all items) to identify the defective items. A test is positive if the pool contains at least one defective items and negative if it contains no defective item. Group testing can be viewed as a sparse inference problem with a combinatorial flavor with diverse applications, for example, clone screening (Balding and Torney 1997), multichannel access in high speed computer networks (Bar-Noy et al. 1992), medical examination (Robert 1943), statistics (Hu, Hwang, and Wang 1981), compressed sensing and machine learning (Atia and Saligrama 2012).

The group testing problem can be separated into two parts: *design* and *decoding* (Aldridge, Johnson, and Scarlett 2019). The design part concerns how to choose the testing strategy, that is, which items to place in which pools. The decoding problem consists of determining which items are defective given the set of tests and their outcomes, and can

be formulated as a combinatorial optimization problem (Du, Hwang, and Hwang 2000; Malioutov and Malyutov 2012; Aldridge, Baldassini, and Johnson 2014). The design phase can be modeled in two ways: adaptive and non-adaptive testing (Aldridge, Johnson, and Scarlett 2019). In adaptive testing, the test pools are designed sequentially, wherein each test depends on the outcome of the previous ones. In the case of non-adaptive testing, all test pools are designed in advance, which allows parallel implementation. In a more practical setting, a noisy variant to group testing is considered where tests are inverted/flipped according to some specific random model or in an adversarial manner (Aldridge, Johnson, and Scarlett 2019). In this work, we pursue the direction of solving the decoding phase in non-adaptive group testing in both noiseless and noisy settings by introducing a novel MaxSAT-based approach.

The primary contribution of this paper is a framework MGT[1] (**M**axSAT-based framework for **G**roup **T**esting), which provides a novel and efficient MaxSAT-based encoding to address the decoding phase of non-adaptive group testing. We consider both noiseless and noisy testing to make the proposed approach applicable in the realistic settings. For the noisy setting, we propose a compact encoding and prove its soundness. Our extensive experimental results show that MGT can solve group testing instances of 10000 items with 3% defectivity, which no prior work can handle to the best of our knowledge. Moreover, the accuracy of MGT is shown better than that of the existing state-of-the-art approaches. In our experiments, we also observe a phase transition between the unrecoverability-recoverability of the input items in terms of computation time. This phase transition in MGT shows a promising connection between two sub fields: group testing and the MaxSAT problem, and demonstrates the practical application of MaxSAT.

The rest of the paper is organized as follows. We discuss related works in Section 2, define notations and preliminaries in Section 3 and formulate the problem in Section 4. In Section 5, we introduce our framework MGT, show the experimental results in Section 6 and conclude in Section 7.

---

[1]https://github.com/meelgroup/mgt

2 Related Works

Robert Dorfman formalized group testing during World War II with the purpose of creating a model to pool sick soldiers without direct testing of every single candidate (Robert 1943). Thus, group testing can be viewed as a pooling strategic problem (Zhang et al. 2013) with the goal of designing an optimal set of tests of items efficiently such that the test results contain enough information to determine a small subset of defective items. Another body of work focuses on the decoding of the pooling results. In particular, Chan et al. (Chan et al. 2011; 2014) find a similarity between compressing sensing (Aksoylar, Atia, and Saligrama 2016; Malyutov 2013) and group testing, and present the COMP (combinatorial orthogonal matching pursuit) algorithm for the decoding phase of non-adaptive group testing in both noiseless and noisy settings. Aldridge et al. (Aldridge, Baldassini, and Johnson 2014) consider non-adaptive noiseless group testing and propose two algorithms: DD (definite defectives) and SCOMP (sequential COMP), which require stronger evidence to declare an item defective; and an essentially optimal but computationally difficult SSS (smallest satisfying set) algorithm. The recent works (Scarlett and Cevher 2016; Coja-Oghlan et al. 2019) have established performance guarantees for similar computationally expensive algorithms. In addition, linear programming (LP) algorithms have been proposed to approximate the SSS algorithm with practical runtime (Malioutov and Malyutov 2012; Malyutov and Sadaka 2009; Chan et al. 2014).

The maximum satisfiability (MaxSAT) problem is an optimization analogue to the SAT (satisfiability) problem. See Section 3 for the detailed definition. MaxSAT is complete for the class FP^{NP}, which includes many practical optimization problems. The added scalability and improvement of MaxSAT solvers in recent years have encouraged researchers to reduce several optimization problems into MaxSAT, for example, optimal planning (Robinson et al. 2010; Zhang and Bacchus 2012), interpretable rule-based classifications in machine learning (Ghosh and Meel 2019; Malioutov and Meel 2018), automotive configuration (Walter, Zengler, and Küchlin 2013), data analysis and machine learning (Berg, Hyttinen, and Järvisalo 2015), automatic test pattern generation for cancer therapy (Lin and Khatri 2012), etc. However, to the best of our knowledge, there is no prior work on group testing that takes benefit from MaxSAT-based solution approach.

3 Notation and Preliminaries

We use capital boldface letters such as $\mathbf{X}$ to denote matrices, while lower boldface letters $\mathbf{x}$ are reserved for vectors/sets. For a matrix $\mathbf{X}$, $\mathbf{X}_i$ represents the i-th row of $\mathbf{X}$ while for a vector/set $\mathbf{x}$, x_i represents the i-th element of $\mathbf{x}$. A Boolean vector $\mathbf{x} \in \{0,1\}^n$ with dimension n is called a *sparse* vector if the number of non-zero elements $\sum_{i=1}^n x_i \ll n$.

3.1 Boolean Logic and CNF

A propositional formula F in Conjunctive Normal Form (CNF) with n boolean variables $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ is defined as a conjunction of clauses, where each clause is a
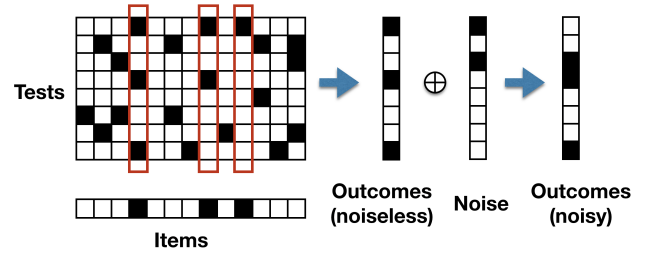

Figure 1: A simple group testing example. Black and white cells represent 1s and 0s, respectively. The test matrix indicates which items are included in the test, and the noiseless test outcome is the bit-wise OR of the columns corresponding to defective items. In the noisy setting, some of the outcomes are flipped.

disjunction of literals. A literal can be either a variable x_i or its complement $\neg x_i$. If σ is an assignment to variables and $x_i \in \mathbf{x}$, we use $\sigma(x_i)$ to denote the value assigned to x_i in σ. The propositional satisfiability (SAT) problem finds a *satisfying assignment* or *witness* σ^* to variables in $\mathbf{x}$ that makes F evaluate to 1 (*true*). Formally, given a CNF formula $F = \bigwedge_i C_i$, $\sigma^* \models F$ iff $\forall i, \sigma^* \models C_i$, wherein $\sigma^* \models C_i$ iff $\exists x \in C_i, \sigma^*(x) = 1$.

In this work, we focus on the weighted variant of CNF wherein a weight function is defined over clauses. For a clause C_i and weight function $\mathsf{wt}(\cdot)$, we use $\mathsf{wt}(C_i)$ to denote the weight of clause C_i. We say that a clause C_i is *hard* if $\mathsf{wt}(C_i) = \infty$; otherwise, C_i is called a *soft* clause. To avoid notational clutter, we overload $\mathsf{wt}(\cdot)$ to denote the weight of an assignment or clause, depending on the context. We define the weight of an assignment σ as the sum of weights of the clauses that σ satisfies. Let $\mathbb{1}\{true\} = 1$ and $\mathbb{1}\{false\} = 0$. Formally, $\mathsf{wt}(\sigma) = \sum_i \mathsf{wt}(C_i) \cdot \mathbb{1}\{\sigma \models C_i\}$.

3.2 MaxSAT

Given a formula F and weight function $\mathsf{wt}(\cdot)$, the problem of MaxSAT is to find an assignment σ^* that has the maximum weight, i.e., $\sigma^* = \mathsf{MaxSAT}(F, \mathsf{wt}(\cdot))$ if $\forall \sigma \neq \sigma^*, \mathsf{wt}(\sigma*) \geq \mathsf{wt}(\sigma)$. Our formulation will have positive clause weights, hence MaxSAT corresponds to satisfying as many clauses as possible, and picking the strongest clauses among the unsatisfied ones. Borrowing terminology from the community focused on developing MaxSAT solvers, we are solving a partial weighted MaxSAT instance wherein we mark all the clauses with ∞ weight as hard, and clauses with other positive value less than ∞ weight as soft, and asking for a solution that optimizes the partial weighted MaxSAT formula. Knowledge of the inner workings of MaxSAT solvers and the encoding of representations into weighted MaxSAT instances are not required for this paper.

3.3 Group Testing Model

Let $\mathbf{x} \in \{0,1\}^n$ be a vector of n items, where $x_i = 1$ denotes a defective item and $x_i = 0$ denotes a non-defective item. Let $k(\mathbf{x}) = \sum_{i=1}^n x_i$ be the number of defective items in $\mathbf{x}$. We will use k instead of $k(\mathbf{x})$ to denote the number of

defective items when it is clear from the context. Consider a matrix of pooled measurements $\mathbf{A} \in \{0,1\}^{m \times n}$ with m tests, where $\mathbf{A}_{i,j}$ denotes the j-th item of the i-th row (or test) of $\mathbf{A}$. Specifically, $\mathbf{A}_{i,j} = 1$ if the j-th item belongs to the i-th test, and $\mathbf{A}_{i,j} = 0$ otherwise. Additionally, we consider the Boolean vector $\mathbf{y} \in \{0,1\}^m$, where y_i represents the outcome of the i-th test:

$$y_i = \begin{cases} 1 & \text{if } \exists j \in \{1, \ldots, n\}, \ \mathbf{A}_{i,j} \wedge x_j = 1 \\ 0 & \text{otherwise} \end{cases}$$

In order to model a noisy setting, we allow tests to return inverted/flipped outcomes. Hence, we consider a Boolean vector $\boldsymbol{\xi} \in \{0,1\}^m$ such that the i-th test gives an inverted outcome iff $\xi_i = 1$ and $\xi_i = 0$ otherwise. In the noisy setting, we modify the definition of $\mathbf{y}$ as follows:

$$y_i = \begin{cases} 1 & \text{if } \exists j \in \{1, \ldots, n\}, \ (\mathbf{A}_{i,j} \wedge x_j) \oplus \xi_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

Here $\oplus$ represents the logical XOR operator. We illustrate a simple example of the group testing problem in Figure 1 (see the caption for a description), showing both the noiseless and noisy outcomes.

We consider Bernoulli trials (Aldridge, Johnson, and Scarlett 2019) to model a stochastic group testing instance. In particular, we consider three independent and identically distributed (i.i.d.) Bernoulli processes: an item is defective independently with probability $p < 0.5$, item i belongs to test j independently with probability q, and in the noisy setting, a test outcome is inverted independently with probability $d < 0.5$.

4 Problem Formulation

Given the measurement matrix $\mathbf{A} \in \{0,1\}^{m \times n}$ and test vector $\mathbf{y} \in \{0,1\}^m$, we attempt to find the smallest set of defective items in $\mathbf{x}$, where each item in $\mathbf{x}$ is defective with equal probability. Our assumption of defectivity being rare, is motivated by compressed sensing[2], which is also applied in the SSS (smallest satisfying set) algorithm (Aldridge, Baldassini, and Johnson 2014) and the LP-relaxation approach (Malioutov and Malyutov 2012). Formally, in the noiseless setting, we minimize the following function subject to the constraints that there is at least one defective item in a positive test and no defective item in a negative test.

$$\min \sum_{j=1}^n x_j \qquad (1)$$

In the noisy setting, each test may be inverted/flipped in one of two ways: original test finds a defective item(s) but the noisy output indicates that all items are non-defective, and vice versa. In the noisy setting, we minimize noise variable ξ_i while preferring $\mathbf{x}$ to be the sparsest (Malioutov and

---

[2]In compressed sensing, one typically seeks the sparsest signal $\mathbf{x}$ that fits some given measurements $\mathbf{y} = \mathbf{A}\mathbf{x}$ (Aldridge, Baldassini, and Johnson 2014).

Malyutov 2012) with i.i.d. prior on both $\mathbf{x}$ and $\boldsymbol{\xi}$. In this setting, the constraints are similar to the noiseless setting except that now a test may be flipped.

$$\min \sum_{j=1}^n x_j + \lambda \sum_{i=1}^m \xi_i \qquad (2)$$

The objective function is formulated as to obtain the *sparsest* and *less noisy* solution, where the parameter $\lambda \in \mathbb{R}^+$ balances the trade-off between the amount of noise and the sparsity of the solution. In Section 5.3, we discuss how to set the value of λ.

5 MGT: A MaxSAT-based Framework for Group Testing

In this section, we describe the primary contribution of this paper, MGT, a MaxSAT-based framework for solving the decoding phase of non-adaptive group testing. We first, in Section 5.1, discuss the MaxSAT encoding for both noiseless and noisy setting. Later, in Section 5.2, we propose a compact MaxSAT encoding for the noisy setting and prove its soundness. Finally, in Section 5.3, we discuss how to set the value of λ.

5.1 MaxSAT Encoding

We first describe the MaxSAT encoding for the noiseless setting and later extend the formulation to noisy setting.

Noiseless setting: We consider a unit soft clause that tries to falsify each x_j in the objective function. The weight of the soft clause is 1, which is derived from the coefficient of x_j in the objective function in Eq. 1:

$$S_j := \neg x_j; \qquad \mathsf{wt}(S_j) = 1.$$

To encode the constraints associated with the tests, we construct hard clauses in the MaxSAT query. Recall that $\mathbf{A}_{i,j} = 1$ denotes the j-th item being included in the i-th test. When the outcome of the i-th test is positive ($y_i = 1$), there must be at least one defective item included in that test. Therefore, we construct the following hard clause when $y_i = 1$:

$$C_i := \bigvee_{j | \mathbf{A}_{i,j}=1} x_j; \qquad \mathsf{wt}(C_i) = \infty.$$

On the other hand, when $y_i = 0$, all items included in the i-th test must be non-defective. Therefore we construct the following hard clause when $y_i = 0$:

$$\widetilde{C}_i := \neg (\bigvee_{j | \mathbf{A}_{i,j}=1} x_j); \qquad \mathsf{wt}\left(\widetilde{C}_i\right) = \infty.$$

In this case, we apply de Morgan's law on clause C_i to convert it into CNF, $\widetilde{C}_i = \bigwedge_{j | \mathbf{A}_{i,j}=1} \widetilde{C}_{ij}$, where $\widetilde{C}_{ij}$ is defined as follows:

$$\widetilde{C}_{ij} := \neg x_j; \qquad \mathsf{wt}\left(\widetilde{C}_{ij}\right) = \infty.$$

Finally, the MaxSAT formula F is the conjunction of all clauses:

$$F := \bigwedge_{j=1}^{n} S_j \wedge \bigwedge_{i=1, y_i=1}^{m} C_i \wedge \bigwedge_{i=1, y_i=0}^{m} \widetilde{C}_i. \qquad (3)$$

Since all clauses $S_j, C_i, \widetilde{C}_{ij}$ are in CNF, no further translation is required. Once the MaxSAT formula F is formulated, an off-the-shelf MaxSAT solver takes formula F and weight $\mathsf{wt}(\cdot)$ as inputs and returns an optimal assignment to the variable x_j. We find the defective items according to the optimal assignment as follows.

Construction 1. *Consider the assignment* $\sigma^* = \mathsf{MaxSAT}(F, \mathsf{wt}(\cdot))$. *Then item j is detected to be defective if* $\sigma^*(x_j) = 1$.

Noisy Setting: In addition to S_j, we construct unit soft clauses that try to falsify each noise variable ξ_i. The weight of each soft clause is λ, which is equal to the coefficient of ξ_i in Eq. 2:

$$N_i := \neg \xi_i; \qquad \mathsf{wt}(N_i) = \lambda.$$

In the noisy setting, test y_i is inverted when $\xi_i = 1$, and otherwise y_i remains same. Hence, we construct hard clauses C_i for positive tests and $\widetilde{C}_i$ for negative tests as follows:

$$C_i := \Big(\bigvee_{j|\mathbf{A}_{i,j}=1} x_j \Big) \oplus \xi_i \qquad \text{when } y_i = 1,$$

$$\widetilde{C}_i := \Big(\bigwedge_{j|\mathbf{A}_{i,j}=1} \neg x_j \Big) \oplus \xi_i \qquad \text{when } y_i = 0.$$

Here the hard clauses C_i and $\widetilde{C}_i$ have XOR operators in their definitions; that is, these clauses are not in CNF. We first translate these clauses into CNF and then call an off-the-shelf MaxSAT solver for the optimal solution. All translated CNF clauses have weight ∞. Next, we discuss the translation in detail.

- For a positive test output ($y_i = 1$), if there are t_i literals in the clause $\bigvee_{j|\mathbf{A}_{i,j}=1} x_j$, standard CNF translation generates $(t_i + 1)$ CNF clauses while translating $(\bigvee_{j|\mathbf{A}_{i,j}=1} x_j) \oplus \xi_i$. For example, $(x_1 \vee x_2 \vee x_3) \oplus \xi$ is translated into 4-clause CNF as $(\neg \xi \vee \neg x_1) \wedge (\neg \xi \vee \neg x_2) \wedge (\neg \xi \vee \neg x_3) \wedge (\xi \vee x_1 \vee x_2 \vee x_3)$.

- For a negative test output ($y_i = 0$), if there are t_i literals in $\bigwedge_{j|\mathbf{A}_{i,j}=1} \neg x_j$, standard CNF translation generates $(t_i + 1)$ CNF clauses while translating $(\bigwedge_{j|\mathbf{A}_{i,j}=1} \neg x_j) \oplus \xi_i$. For example, $(\neg x_1 \wedge \neg x_2 \wedge \neg x_3) \oplus \xi$ is translated into 4-clause CNF as $(\neg x_1 \vee \xi) \wedge (\neg x_2 \vee \xi) \wedge (\neg x_3 \vee \xi) \wedge (\neg \xi \vee x_1 \vee x_2 \vee x_3)$.

Once all clauses are translated into CNF, we formulate F as follows.

$$F := \bigwedge_{j=1}^{n} S_j \wedge \bigwedge_{i=1}^{m} N_i \wedge \bigwedge_{i=1, y_i=1}^{m} C_i \wedge \bigwedge_{i=1, y_i=0}^{m} \widetilde{C}_i \qquad (4)$$

In the noisy setting, we find the defective items and noisy tests from the optimal assignment to the variables of F as follows.

Construction 2. *Consider the assignment* $\sigma^* = \mathsf{MaxSAT}(F, \mathsf{wt}(\cdot))$. *Then item j is detected to be defective if* $\sigma^*(x_j) = 1$, *and test i is declared inverted if* $\sigma^*(\xi_i) = 1$.

5.2 A Compact Encoding for Noisy Setting:

Since the MaxSAT query in the noisy setting has XOR operators in the definition, its translation to CNF generates additional clauses. We now discuss on how to reduce the number of clauses in the encoding with the aim of reducing the runtime. The improved encoding leverages the soft clauses in Eq. 4. Formally, we propose a compact encoding, where we replace XOR with OR in both C_i and $\widetilde{C}_i$ and define relaxed clauses C_i' and $\widetilde{C}_i'$ respectively. We first define the relaxed clauses C_i' and $\widetilde{C}_i'$, and then provide the theoretical guarantee of the optimal solution of the compact encoding:

$$C_i' := \Big(\bigvee_{j|\mathbf{A}_{i,j}=1} x_j \Big) \vee \xi_i \qquad \text{when } y_i = 1,$$

$$\widetilde{C}_i' := \Big(\bigwedge_{j|\mathbf{A}_{i,j}=1} \neg x_j \Big) \vee \xi_i \qquad \text{when } y_i = 0.$$

Let $\sigma = \mathsf{MaxSAT}(F, \mathsf{wt}(\cdot))$ be the optimal assignment to the variables in the MaxSAT formula F. Here we slightly abuse notation and define $\sigma \underset{\text{opt}}{\models} F$ to represent that σ is the optimal solution of a MaxSAT formula F.

Theorem 1. *Let F be a MaxSAT formula in the noisy setting, and F' be the compact encoded formula with the above relaxation of XOR. Then σ is an optimal solution to F iff σ is an optimal solution to F', i.e., although $F \neq F'$, it holds that* $\sigma \underset{\text{opt}}{\models} F \Leftrightarrow \sigma \underset{\text{opt}}{\models} F'$.

The intuition behind Theorem 1 is that an optimal solution to F minimizes a weighted sum of the sparsity and the number of tests flipped, and an optimal solution to F' minimizes a weighted sum of the sparsity and the number of tests whose constraints are 'ignored'. By taking the tests marked 'flipped' in F and marking them as 'ignored' in F', we find that any solution to F has a matching solution to F'. Moreover, the optimal solution to F' does not label 'ignored' for any test that is already consistent, since any solution doing so could be improved by marking that test as 'not ignored'. Hence, in any optimal solution to F', replacing 'ignored' by 'flipped' gets us to a matching solution to F.

We proceed by formalizing this intuition.

Proof. Since F and F' differ only in the hard clauses C_i and C_i' (similarly $\widetilde{C}_i$ and $\widetilde{C}_i'$), and an optimal assignment of a MaxSAT formula always satisfies all hard clauses, the result $\sigma \underset{\text{opt}}{\models} F' \Leftrightarrow \sigma \underset{\text{opt}}{\models} F$ will follow once we prove that

$$\Big(\sigma \underset{\text{opt}}{\models} F' \Rightarrow \sigma \models C_i\Big) \wedge \Big(\sigma \underset{\text{opt}}{\models} F \Rightarrow \sigma \models C_i'\Big) \qquad (5)$$

for all tests with positive outcomes, and

$$(\sigma \underset{\text{opt}}{\models} F' \Rightarrow \sigma \models \widetilde{C}_i) \wedge (\sigma \underset{\text{opt}}{\models} F \Rightarrow \sigma \models \widetilde{C}'_i) \quad (6)$$

for all tests with negative outcomes.

To prove (5), we first show that $\sigma \underset{\text{opt}}{\models} F' \Rightarrow \sigma \models C_i$. Expanding the definition of $C'_i := (\bigvee_{j|\mathbf{A}_{i,j}=1} x_j) \vee \xi_i$, if ξ_i is assigned false in the optimal solution, i.e., $\sigma \models \neg \xi_i$ and $\sigma \models C'_i$, then $\sigma \models C_i$ because $\sigma \models \bigvee_{j|\mathbf{A}_{i,j}=1} x_j$. We proceed by considering the other case, namely, $\sigma \models \xi_i$.

Recall that both F and F' have soft clauses $S_j := \neg x_j$ and $N_i := \neg \xi_i$. When $\sigma \models \xi_i$ and $\sigma \models C'_i$, the optimal assignment σ for F' cannot satisfy both ξ_i and $\bigvee_{j|\mathbf{A}_{i,j}=1} x_j$ simultaneously: If both were satisfied, the same assignment with ξ_i changed to zero would be feasible and attain a smaller overall weight. Combining this observation with the case $\sigma \models \neg \xi_i$ above, we deduce that any test with $y_i = 1$ must satisfy $\sigma \models (\bigvee_{j|\mathbf{A}_{i,j}=1} x_j) \oplus \xi_i$, i.e., $\sigma \models C_i$.

To prove that $\sigma \underset{\text{opt}}{\models} F \Rightarrow \sigma \models C'_i$, observe that the hard clause $\sigma \models C_i$ implies that either $\sigma \models \xi_i$ or $\sigma \models (\bigvee_{j|\mathbf{A}_{i,j}=1} x_j)$. Combining these gives $\sigma \models (\bigvee_{j|\mathbf{A}_{i,j}=1} x_j) \vee \xi_i$, i.e., $\sigma \models C'_i$.

Similar reasoning applies to (6) (i.e., the negative tests), and we deduce that $\sigma \underset{\text{opt}}{\models} F' \Leftrightarrow \sigma \underset{\text{opt}}{\models} F$.

□

5.3 Deciding the Value of λ

In the objective function for the noisy setting in Eq. 2, the parameter λ decides the trade-off between the number of defective items and the number of inverted tests. Since we model both defectivity and noise using i.i.d. Bernoulli trials in the stochastic model, we can set the appropriate value of λ from the associated probabilities p and d.

Suppose each item is defective independently with probability $p < 0.5$ and each test gets inverted independently with probability $d < 0.5$. Then for a candidate defective set $\mathbf{x}$ and recovered set $\widehat{\mathbf{x}}$, finding the optimal solution requires minimizing $\Pr[\widehat{\mathbf{x}} \neq \mathbf{x}]$, which is equivalent to maximizing the posterior probability $\Pr[\mathbf{x}|\mathbf{A}, \mathbf{y}]$ (Aldridge, Baldassini, and Johnson 2014). Now we derive $\Pr[\mathbf{x}|\mathbf{A}, \mathbf{y}]$ as follows:

$$\Pr[\mathbf{x}|\mathbf{A}, \mathbf{y}] = \frac{\Pr[\mathbf{x}, \mathbf{A}, \mathbf{y}]}{\Pr[\mathbf{A}, \mathbf{y}]} \propto \Pr[\mathbf{x}]\Pr[\mathbf{y}]$$
$$= p^{k(\mathbf{x})}(1-p)^{n-k(\mathbf{x})}d^{\tau(\mathbf{x},\mathbf{A},\mathbf{y})}(1-d)^{m-\tau(\mathbf{x},\mathbf{A},\mathbf{y})}$$
$$= (1-p)^n \left(\frac{p}{1-p}\right)^{k(\mathbf{x})}(1-d)^m \left(\frac{d}{1-d}\right)^{\tau(\mathbf{x},\mathbf{A},\mathbf{y})}$$
$$\Rightarrow \Pr[\mathbf{x}|\mathbf{A}, \mathbf{y}] \propto \left(\frac{p}{1-p}\right)^{k(\mathbf{x})}\left(\frac{d}{1-d}\right)^{\tau(\mathbf{x},\mathbf{A},\mathbf{y})}$$

Here $k(\mathbf{x})$ is the number of defective items, and $\tau(\mathbf{x}, \mathbf{A}, \mathbf{y})$ is the Hamming distance between the output $\mathbf{y}$ and the expected output that one would get if there were no noise (that is, the number of inverted tests).

Taking the log of this posterior, then normalizing and equating with Eq. 2, we find that $\lambda = (\log \frac{d}{1-d})/(\log \frac{p}{1-p})$ is the optimal choice.

6 Experimental Results

We have developed a prototype implementation of MGT to solve the decoding phase of non-adaptive group testing in both the noiseless and noisy settings. In our implementation, we employ MaxHS (Davies and Bacchus 2011) as the underlying MaxSAT solver. We compare MGT with a state-of-the-art approach namely, the approximated linear programming relaxation approach (LP) that also solves the decoding phase of non-adaptive group testing (Malioutov and Malyutov 2012). In the LP approach, we use CPLEX as the underlying LP solver[3]. We set the cut-off time of both the LP and MaxHS solvers to be 100 seconds. If an optimal solution is not found within the cut-off time, both solvers return the current best solution. The experiment was conducted on a machine with Intel core i7 (3.4 GHz) and 8 GB of RAM.

To model a group testing instance, we first choose the number of defective items k for a fixed number of items n ($k \in [0.01n, 0.1n]$). Then the item vector $\mathbf{x}$ is generated with k defective items. Once the number of defective items is known, we consider a Bernoulli process with probability $q = (\log 2)/k$ to construct measurement matrix $\mathbf{A}$, which is motivated by Aldridge at el. (Aldridge, Johnson, and Scarlett 2019). In the noisy setting, we consider another Bernoulli process with probability $d < 0.5$ to generate the noise vector $\boldsymbol{\xi}$. We have used $p = \frac{k}{n}$ to derive λ as described in Section 5.3. We repeat each experiment for $l = 100$ trials to ensure statistical consistency.

6.1 Evaluation Metrics

We evaluate MGT with LP based on the Hamming distance and probability of success, which are defined below.

- **Hamming Distance:** Given an item vector $\mathbf{x}$ and the recovered solution $\widehat{\mathbf{x}}$, the Hamming distance $h(\mathbf{x}, \widehat{\mathbf{x}})$ is the number of items that are wrongly detected in either of the two ways: a non-defective item detected as defective, or a defective item detected as non-defective.

- **Probability of Success s:** The probability of success is defined as the probability of attaining zero Hamming distance in the recovered solution. Formally, $s = \frac{\sum_{i=1}^{l} \mathbb{1}\{h(\mathbf{x}_i,\widehat{\mathbf{x}}_i)=0\}}{l}$, where $\mathbf{x}_i$ is the randomly generated item vector at the i-th trial and $\widehat{\mathbf{x}}_i$ is the associated recovered solution.

The primary objective of our study is to seek answer to the following questions:

1. How does MGT scale with respect to the number of items n, the number of defective items k, and the number of tests m?

2. How do the accuracy and runtime performance of MGT compare to existing state-of-the-art approaches?

---

[3]When the LP solver returns a non-integer value in the solution, we adopt the standard procedure to round to 1 when the value is > 0.5 and 0 otherwise.
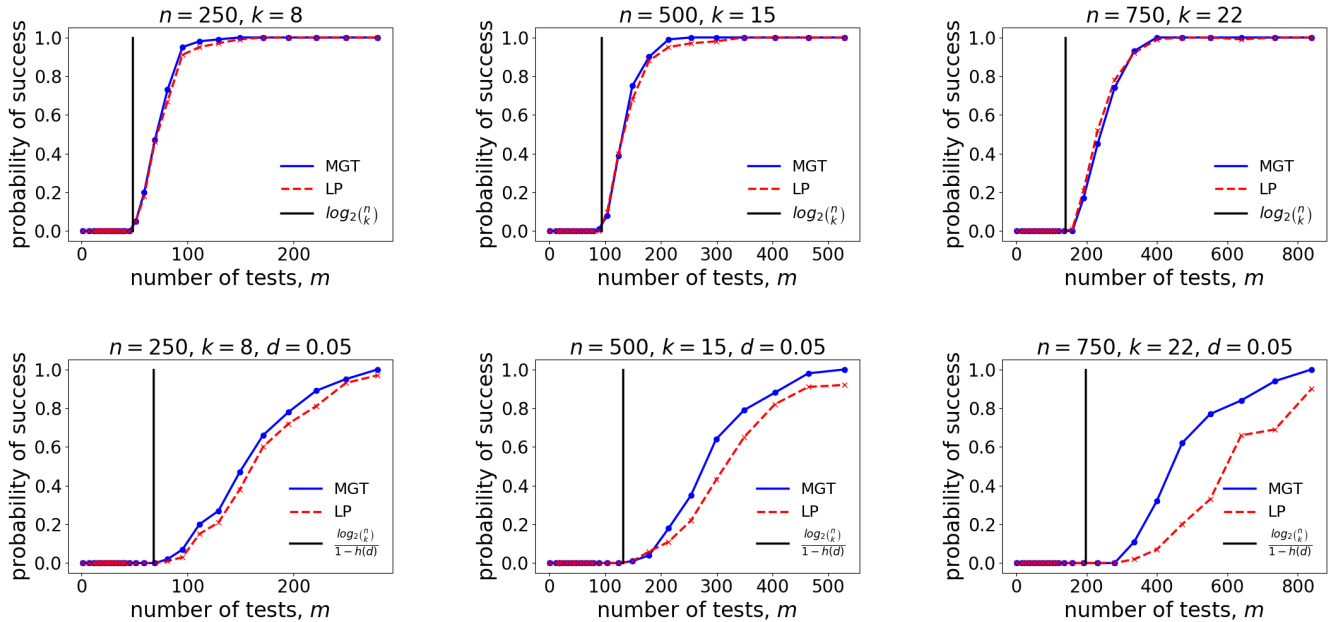
Figure 2: Effect of the number of items n on the probability of success in both the noiseless (figures in the top row) and noisy (figures in the bottom row) settings. In each graph $k = 0.03n$, and for the noisy setting, $d = 0.05$.

3. Do we observe any interesting behavior of the runtime of MGT?

4. Does MGT follow known bounds on the number of tests for recovery?

5. How efficient is the compact encoding compared to the naive encoding in the noisy setting?

To summarize our empirical study, MGT shows impressive scalability and can solve group testing instances with at least $n = 10000$ and $k = 0.03n$ in the noiseless setting, which no prior work can handle to the best of our knowledge. Moreover, the accuracy of MGT is better than the compared approach. In addition, the runtime behavior of MGT in the noiseless setting indicates a phase transition revealing the easy-hard-easy nature of group testing. We also find that MGT follows known bounds on the number of tests for recovery. Finally, in the noisy setting, the compact encoding outperforms the naive encoding w.r.t. runtime by a large margin, which highlights the practical applicability of the compact encoding with added optimality guarantee. In the following, we provide the details of our experiments.

6.2 Empirical Results:

Varying the Number of Items: We first observe whether MGT follows the theoretical bound on the number of tests m for non-zero probability of success ($s > 0$) in both the noiseless and noisy settings. In Figure 2, we show graphs where we vary m and plot the corresponding s for different choices of the number of items n. For each choice of n, we set the number of defective items to $k = 0.03n$. The black line in each graph denotes the bound on m, which

is $\log_2 \binom{n}{k}$ for the noiseless setting (Baldassini, Johnson, and Aldridge 2013; Aldridge, Johnson, and Scarlett 2019) and $\log_2 \binom{n}{k}/(1 - h(d))$ for the noisy setting (Scarlett 2018; Johnson 2017), where d is the probability on noisy tests, and $h(d) = -d \log_2 d - (1-d) \log_2(1-d)$ is the binary entropy in bits. In all graphs, we find that empirically s becomes non-zero after m exceeds the theoretical bound in both MGT and LP.

Moreover, s increases and becomes closer to 1 as m increases and becomes closer to n. As we observe more closely, we find that in the noiseless setting, s quickly becomes 1, whereas more tests are required to reach the same level of s when we consider the noisy setting. In terms of accuracy, MGT and LP show a similar performance in the noiseless setting. However, MGT outperforms LP in the noisy setting, and the difference is greater for higher values of n. This result suggests the effectiveness of MGT over LP in terms of the quality of the recovered solutions.

Varying the Number of Defective Items: In Figure 3, we vary the number of defective items $k \in [0.01n, 0.1n]$ and show its effect on the probability of success in both the noiseless and the noisy settings. Empirically we find that for a fixed number of items n, as we increase the number of defective items k, more tests are required to reach the same level of the probability of success in both MGT and LP. In addition, when k is much higher ($\approx 0.1n$), near-perfect recovery often becomes unachievable for random Bernoulli tests even if $m = n$, in both the noiseless and noisy settings.

Scalability Analysis: We experiment with higher values of the number of items n to observe the scalability perfor-
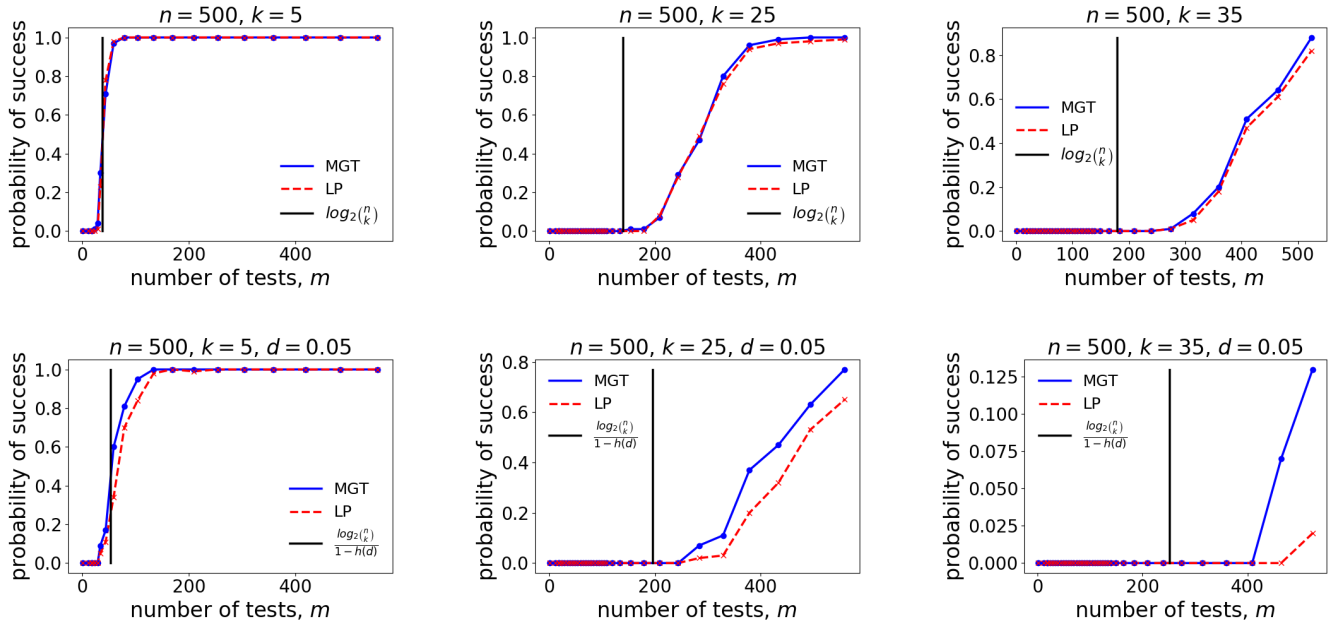
Figure 3: Effect of the number of defective items k on the probability of success in both the noiseless (figures in the top row) and noisy (figures in the bottom row) settings. In each graph, $n = 500$ and for the noisy setting, $d = 0.05$.

mance of both MGT and LP in the noiseless setting, and show the result in Figure 4. Both LP and MaxSAT solvers are given equal computation resources, but LP shows a memory limit when $n > 2000$. Interestingly, we find that MGT shows impressive scalability for higher values of n. In particular, we run MGT for $n = 10000$ and $k = 0.03n$ successfully, and it can potentially go further.

We have also experimented with higher defective items for $n = 10000$, where the accuracy of MGT starts deteriorating even if we allow more tests. To conclude, this scala-

bility result shows the promise of applying MaxSAT solvers in various practical combinatorial optimization problems.

Runtime Analysis and Phase Transition: We compute the average runtime of all trials, and in Figure 5, we present it while varying the number of tests m for both the noiseless and the noisy settings. In all graphs, we find that both MGT and LP require more runtime in the noisy setting than in the noiseless setting, which points out the effect of noise on the computation process. Moreover, LP takes less time than
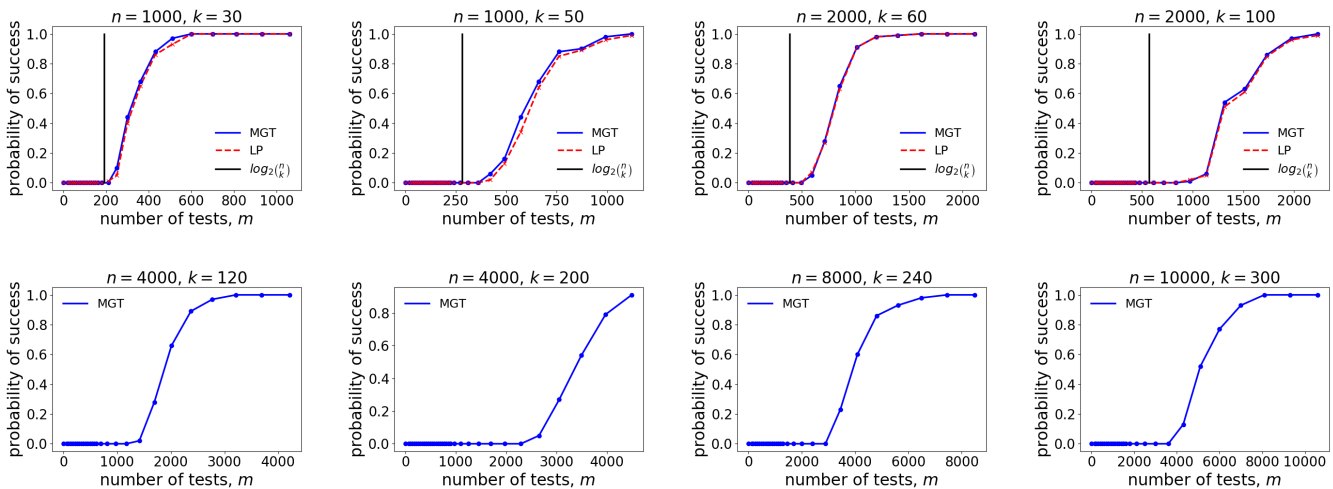

Figure 4: Probability of success for increasing values of n in the noiseless setting.
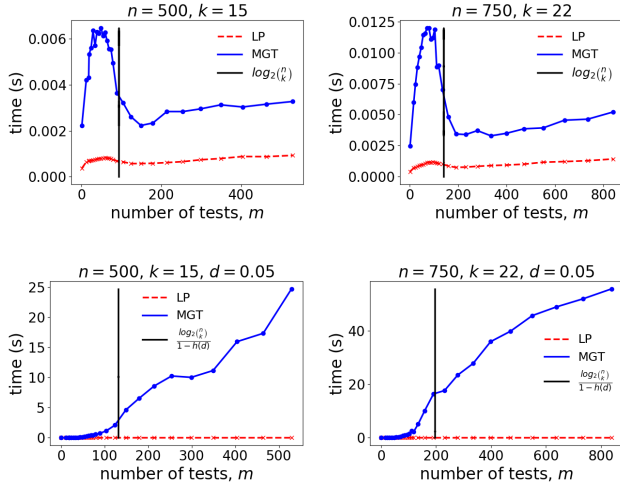
Figure 5: Computation time in the decoding phase of group testing in both the noiseless (figures in the top row) and noisy (figures in the bottom row) settings. In each graph, $k = 0.03n$, and for the noisy setting $d = 0.05$.

Figure 6: Runtime performance of the compact encoding and the naive encoding for the noisy setting of non-adaptive group testing.

MGT as m increases in both settings, and the difference is greater in the noisy setting.

In this context, LP applies relaxation in the integer programs to achieve performance boost in terms of runtime and finds an approximate solution. However, in this MaxSAT-based formulation, the MaxSAT solver aims to find an optimal solution while costing relatively higher runtime.

In the noiseless setting, we find a phase transition of runtime in both MGT and LP. Specifically, the runtime increases until reaching a peak near the theoretical bound on m, then decreases quickly, and again increases as m increases. This observation indicates the easy-hard-easy nature of group testing, where the computational difficulty of a problem instance changes due to changing the number of tests m. Moreover, the theoretical bound on tests denotes that the probability of success also goes from zero to nonzero during this transition. This observation is similar to the phase transition observed in the SAT problem of random k-CNF formulas (Coja-Oghlan and Panagiotou 2013; Chvátal and Reed 1992; Ding, Sly, and Sun 2015). A random k-CNF formula (each clause in the formula has at most k literals) goes from SAT to UNSAT as the number of clauses increases, and the computation time near the transition gives a similar trend to what we have observed here. Although such phase transition is known in SAT instances, in our study we find that the decoding phase of noiseless non-adaptive group testing shows similar characteristics. In particular, the number of tests in group testing can be related to the number of clauses in k-CNF because each test constructs a clause (when test is positive) or a set of clauses (when test is negative), and the SAT-UNSAT transition can be related to the unrecoverability-recoverability of group testing. In future, we will investigate more on the potential bridge between group testing and MaxSAT and such study will be useful
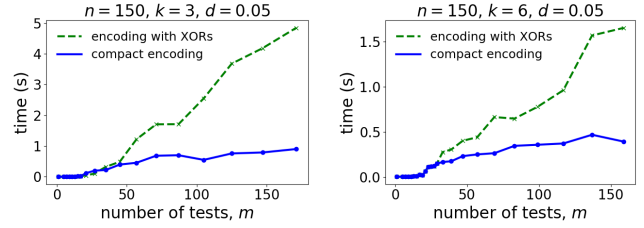
for the improvement of MaxSAT solving.

Efficiency of the Compact Encoding: In Figure 6, we show the runtime performance of both the compact encoding and the naive encoding (encoding with XORs) on the similar choice of parameters. We empirically find that the runtime of the naive encoding is significantly higher than that of the compact encoding as we consider more tests. Moreover, for higher values of n and k, the naive encoding often becomes intractable.

This result suggests that the compact encoding not only provides a theoretical guarantee of optimality, but is also efficient in practice than the naive encoding.

7   Conclusions

In this paper, we presented MGT: a novel MaxSAT-based formulation for solving the decoding phase of non-adaptive group testing in both the noiseless and the noisy settings. For the noisy setting, we proposed a compact encoding with a proof of soundness and empirically showed its efficiency w.r.t. runtime. Our extensive experimental results suggest that MGT shows better accuracy than the compared approach. In addition, MGT shows impressive scalability and solves group testing instances with 10000 items and 3% defective items and possibly beyond. We also observed a phase transition in the runtime of noiseless group testing, which exposes a potential bridge between group testing and MaxSAT. An interesting direction of future work would be to extend our techniques to adaptive setting.

Acknowledgments

References

Aksoylar, C.; Atia, G. K.; and Saligrama, V. 2016. Sparse signal processing with linear and nonlinear observations: A unified shannon-theoretic approach. *IEEE Transactions on Information Theory* 63(2):749–776.

Aldridge, M.; Baldassini, L.; and Johnson, O. 2014. Group testing algorithms: Bounds and simulations. *IEEE Transactions on Information Theory* 60(6):3671–3687.

Aldridge, M.; Johnson, O.; and Scarlett, J. 2019. Group testing: an information theory perspective. *arXiv preprint arXiv:1902.06002*.

Atia, G. K., and Saligrama, V. 2012. Boolean compressed sensing and noisy group testing. *IEEE Transactions on Information Theory* 58(3):1880–1901.

Baldassini, L.; Johnson, O.; and Aldridge, M. 2013. The capacity of adaptive group testing. In *2013 IEEE International Symposium on Information Theory*, 2676–2680. IEEE.

Balding, D. J., and Torney, D. C. 1997. The design of pooling experiments for screening a clone map. *Fungal Genetics and Biology* 21(3):302 – 307.

Bar-Noy, A.; Hwang, F. K.; Kessler, I.; and Kutten, S. 1992. A new competitive algorithm for group testing. In *Proc. IEEE INFOCOM '92: The Conference on Computer Communications*, 786–793 vol.2.

Berg, J.; Hyttinen, A.; and Järvisalo, M. 2015. Applications of MaxSAT in data analysis. *Pragmatics of SAT*.

Chan, C. L.; Che, P. H.; Jaggi, S.; and Saligrama, V. 2011. Non-adaptive probabilistic group testing with noisy measurements: Near-optimal bounds with efficient algorithms. In *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 1832–1839. IEEE.

Chan, C. L.; Jaggi, S.; Saligrama, V.; and Agnihotri, S. 2014. Non-adaptive group testing: Explicit bounds and novel algorithms. *IEEE Transactions on Information Theory* 60(5):3019–3035.

Chvátal, V., and Reed, B. 1992. Mick gets some (the odds are on his side)(satisfiability). In *Proceedings., 33rd Annual Symposium on Foundations of Computer Science*, 620–627. IEEE.

Coja-Oghlan, A., and Panagiotou, K. 2013. Going after the k-SAT threshold. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, 705–714. ACM.

Coja-Oghlan, A.; Gebhard, O.; Hahn-Klimroth, M.; and Loick, P. 2019. Information-theoretic and algorithmic thresholds for group testing. *arXiv preprint arXiv:1902.02202*.

Davies, J., and Bacchus, F. 2011. Solving MaxSAT by solving a sequence of simpler SAT instances. In *International conference on principles and practice of constraint programming*, 225–239. Springer.

Ding, J.; Sly, A.; and Sun, N. 2015. Proof of the satisfiability conjecture for large k. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, 59–68. ACM.

Du, D.; Hwang, F. K.; and Hwang, F. 2000. *Combinatorial group testing and its applications*, volume 12. World Scientific.

Ghosh, B., and Meel, K. S. 2019. IMLI: An incremental framework for MaxSAT-based learning of interpretable classification rules. In *Proceedings of AAAI/ACM Conference on AI, Ethics, and Society(AIES)*.

Hu, M.; Hwang, F.; and Wang, J. K. 1981. A boundary problem for group testing. *SIAM Journal on Algebraic Discrete Methods* 2(2):81–87.

Johnson, O. 2017. Strong converses for group testing from finite blocklength results. *IEEE Transactions on Information Theory* 63(9):5923–5933.

Lin, P.-C. K., and Khatri, S. P. 2012. Application of MaxSAT-based ATPG to optimal cancer therapy design. *BMC genomics* 13(6):S5.

Malioutov, D., and Malyutov, M. 2012. Boolean compressed sensing: LP relaxation for group testing. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3305–3308. IEEE.

Malioutov, D., and Meel, K. S. 2018. MLIC: A MaxSAT-based framework for learning interpretable classification rules. In *International Conference on Principles and Practice of Constraint Programming*, 312–327. Springer.

Malyutov, M., and Sadaka, H. 2009. Capacity of screening under linear programming analysis. In *Proceedings of the 6th Simulation International Workshop on Simulation, St. Petersburg, Russia*, 1041–1045.

Malyutov, M. 2013. Search for sparse active inputs: A review. In *Information Theory, Combinatorics, and Search Theory*, 609–647. Springer.

Robert, D. 1943. The detection of defective members of large populations. *The Annals of Mathematical Statistics* 14:436–440.

Robinson, N.; Gretton, C.; Pham, D. N.; and Sattar, A. 2010. Partial weighted MaxSAT for optimal planning. In *Pacific rim international conference on artificial intelligence*, 231–243. Springer.

Scarlett, J., and Cevher, V. 2016. Phase transitions in group testing. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, 40–53. Society for Industrial and Applied Mathematics.

Scarlett, J. 2018. Noisy adaptive group testing: Bounds and algorithms. *IEEE Transactions on Information Theory* 65(6):3646–3661.

Walter, R.; Zengler, C.; and Küchlin, W. 2013. Applications of MaxSAT in automotive configuration. In *Configuration Workshop*, volume 1, 21. Citeseer.

Zhang, L., and Bacchus, F. 2012. MaxSAT heuristics for cost optimal planning. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*.

Zhang, P.; Krzakala, F.; Mézard, M.; and Zdeborová, L. 2013. Non-adaptive pooling strategies for detection of rare faulty items. In *Proc. IEEE International Conference on Communications Workshops (ICC)*, 1409–1414. IEEE.