arXiv:2006.01325v6 [cs.IT] 29 Jul 2021

Optimal Non-Adaptive Probabilistic Group Testing in General Sparsity Regimes

Wei Heng Bay, Eric Price, and Jonathan Scarlett

Abstract

In this paper, we consider the problem of noiseless non-adaptive probabilistic group testing, in which the goal is high-probability recovery of the defective set. We show that in the case of n items among which k are defective, the smallest possible number of tests equals $\min\{C_{k,n}k\log n, n\}$ up to lower-order asymptotic terms, where $C_{k,n}$ is a uniformly bounded constant (varying depending on the scaling of k with respect to n) with a simple explicit expression. The algorithmic upper bound follows from a minor adaptation of an existing analysis of the Definite Defectives (DD) algorithm, and the algorithm-independent lower bound builds on existing works for the regimes $k \leq n^{1-\Omega(1)}$ and $k = \Theta(n)$. In sufficiently sparse regimes (including $k = o(\frac{n}{\log n})$), our main result generalizes that of Coja-Oghlan *et al.* (2020) by avoiding the assumption $k \leq n^{1-\Omega(1)}$, whereas in sufficiently dense regimes (including $k = \omega(\frac{n}{\log n})$), our main result shows that individual testing is asymptotically optimal for any non-zero target success probability, thus strengthening an existing result of Aldridge (2019) in terms of both the error probability and the assumed scaling of k.

1 Introduction

The group testing problem was originally studied in the context of testing blood samples for rare diseases [13], with the key idea being to reduce the required number of tests via pooling. Group testing has since found applications in communications [16], information retrieval [12], compressed sensing [19], and most recently, COVID-19 testing [30].

The problem is formally defined as follows: There are n items $[n] = \{1, 2, \dots, n\}$, a subset $S \subseteq [n]$ of which is defective, with $|S| = k$. A number of tests are performed, each taking as input a subset of items, and returning positive if and only if the subset contains at least one defective item. A group testing algorithm specifies the number of tests T, the items included in each test, and a decoder that returns an estimate $\hat{S}$ of the defective set given the test outcomes. We are interested in the required number of tests to attain asymptotically vanishing error probability, i.e., $\lim_{n\to\infty} \mathbb{P}[\hat{S} \neq S] = 0$.

We focus on the non-adaptive setting, in which all tests must be specified prior to observing any outcomes; this is often highly desirable in applications, since it permits the tests to be implemented in parallel. In this setting, the tests can be represented as a test matrix $\mathbf{X} \in \{0,1\}^{T\times n}$, where the (i,j)-th entry is 1 if and only if the i-th test contains the j-th item. The test outcomes are then given by the element-wise "OR" of the k columns corresponding to the defective items. Mathematically, the i-th outcome is given by

$$Y^{(i)} = \bigvee_{j\in S} X_j^{(i)}, \tag{1.1}$$

where $X_j^{(i)}$ is the (i,j)-th entry of $\mathbf{X}$.

W.H. Bay and J. Scarlett are with the Department of Computer Science and the Department of Mathematics, National University of Singapore (e-mail: bayweiheng@gmail.com, scarlett@comp.nus.edu.sg). J. Scarlett is also with the Institute of Data Science, National University of Singapore. E. Price is with the Department of Computer Science, University of Texas at Austin (e-mail: ecprice@cs.utexas.edu).

E. Price was supported in part by NSF Award CCF-1751040 (CAREER). J. Scarlett was supported by an NUS Early Career Research Award.

We place a random model on the defective set S. Throughout the majority of the paper, we assume that each item is included in S (i.e., is defective) independently with some probability p (possibly depending on n), referred to as the *prevalence*. We also consider a closely-related model in which k is fixed, and S is a uniformly random subset of $[n]$ of cardinality k. We refer to these models as the *i.i.d. prior* and *combinatorial prior* respectively. The two are closely related, since under the i.i.d. model we have $k = np(1 + o(1))$ with probability approaching one as long as $np = \omega(1)$. See [5, Sec. 1.A] for a more detailed summary of the connections between these models.

Throughout the paper, we make the mild assumption that $p \leq \frac{1}{2}$ (i.i.d. prior) or $k \leq \frac{n}{2}$ (combinatorial prior). Otherwise, the problem fails to be sparse, and it is already well-established that individual (one-by-one) testing is optimal even when adaptivity is allowed [2,25,29]. In addition, our analysis applies essentially unchanged when this factor of $\frac{1}{2}$ is replaced by any fixed constant less than one.

2 Existing Results and Contributions

Here we state the most relevant existing results on probabilistic non-adaptive group testing, and state our own main results in the context of these existing ones. For consistency with the vast majority of existing works, we express the previously-known results in terms of k and n, corresponding to the combinatorial prior. However, the same results apply under the i.i.d. prior when k is replaced by $\overline{k} := np$ throughout, under the mild assumption that $\overline{k} \to \infty$ as $n \to \infty$.

2.1 Main Existing and New Results

A simple counting-based (or entropy-based) argument reveals that the number of tests for the high-probability recovery of S must satisfy $T \geq (1 - o(1)) \log_2 \binom{n}{k}$, or more simply $T = \Omega(k \log \frac{n}{k})$ [6–8, 24]. Moreover, this scaling is order-optimal in the following widely-considered regimes:

- If $k \leq n^{1-\Omega(1)}$, then we have $k \log \frac{n}{k} = \Theta(k \log n)$, and thus, the lower bound matches the ubiquitous $O(k \log n)$ upper bound obtained via random testing [4, 8, 18, 24] or certain explicit designs [21].
- If $k = \Theta(n)$, then we have $k \log \frac{n}{k} = \Theta(n)$, and thus, the lower bound matches the trivial $O(n)$ upper bound corresponding to testing each item individually.

While these observations cover the majority of scaling regimes, there remain "mildly sublinear" regimes in which the existing upper and lower bounds do not match, namely, $k = \Theta(\frac{n}{f(n)})$ for any $f(n)$ satisfying $f(n) = \omega(1)$ and $f(n) = o(n^c)$ for all $c > 0$. A notable example of such a regime is $k = \frac{n}{\text{poly}(\log n)}$. Our first main result, stated below, closes this gap by showing that the correct scaling is always $\Theta(\min\{k \log n, n\})$.

Before stating the result, we introduce the following threshold:

$$T^*(n, k) = \max\left\{k \log_2 \frac{n}{k}, \frac{k \log_2 k}{\ln 2}\right\}. \tag{2.1}$$

While the above discussion focuses on scaling laws, recent refined analyses [1, 10, 11, 22] have nailed down the precise constants in the above-mentioned scaling regimes:

- When $k \leq n^{1-\Omega(1)}$, the optimal threshold for non-adaptive group testing is T^* [11]. Specifically, there exists a strategy using $T \leq (1 + \epsilon)T^*$ tests that succeeds with probability approaching one and has decoding time polynomial in n, whereas any algorithm requires $T \geq (1 - \epsilon)T^*$ to have a success probability bounded away from zero.
- When $k = \Theta(n)$, the optimal threshold for non-adaptive group testing is n [1]. Specifically, with $T = n$ one can trivially use one-by-one testing, whereas any strategy attaining success probability arbitrarily close to one must have $T \geq n - 1$.[1]

[1] The subtraction of one is merely due to the fact that under the combinatorial prior, knowing the status of $n - 1$ items also implies knowing the status of the remaining item.

Based on these results, a reasonable guess is that the optimal threshold for group testing in general scaling regimes is $\min\{T^*(n,k), n\}$ (which scales as $\Theta(\min\{k\log n, n\})$, consistent with the above discussion). Our main result, stated as follows, reveals that this is indeed the case.

Theorem 2.1. *In the non-adaptive group testing problem with n items and prevalence p (possibly depending on n) under the i.i.d. prior with $p = \omega(\frac{1}{n})$ and $p \le \frac{1}{2}$,[2] we have the following for any $\epsilon > 0$:*

- *There exists a test design and polynomial-time decoding algorithm using $T \le \min\{(1+\epsilon)T^*(n, np), n\}$ tests and having a success probability approaching one as $n \to \infty$.*
- *Any group testing strategy having success probability bounded away from zero as $n \to \infty$ must use at least $T \ge (1-\epsilon)\min\{T^*(n, np), n\}$ tests.*

While we find it most convenient to establish this result for the i.i.d. prior, we can use known connections between the two priors to establish the following analog for the combinatorial prior.

Corollary 2.1. *In the non-adaptive group testing problem with n items and $k \le \frac{n}{2}$ defectives under the combinatorial prior, we have the following for any $\epsilon > 0$:*

- *There exists a test design and polynomial-time decoding algorithm using $T \le \min\{(1+\epsilon)T^*(n, k), n\}$ tests and having a success probability approaching one as $n \to \infty$.*
- *Any group testing strategy having success probability bounded away from zero as $n \to \infty$ must use at least $T \ge (1-\epsilon)\min\{T^*(n, k), n\}$ tests.*

These results not only show that optimal non-adaptive group testing requires $T = \Theta(\min\{k\log n, n\})$ tests in general sparsity regimes, but also provide the precise underlying constants.

The algorithmic upper bounds in Theorem 2.1 and Corollary 2.1 are already known in the regime $k \le n^{1-\Omega(1)}$ [11, 22], and follow trivially from one-by-one testing when $\min\{(1+\epsilon)T^*, n\} = n$. Hence, it suffices to establish success using $(1+\epsilon)T^*$ tests in the regime $k = n^{1-o(1)}$. Fortunately, although the analysis of the Definite Defectives (DD) algorithm in [22] was formally only stated for $k \le n^{1-\Omega(1)}$, the analysis can be adapted to the regime $k = n^{1-o(1)}$ with only minor modifications. We detail the required changes in Appendix A.

As for the algorithm-independent lower bounds, with the regime $k \le n^{1-\Omega(1)}$ having been solved in [11], we can again focus on the regime $k = n^{1-o(1)}$ (including $k = \Theta(n)$). In this case, we were unable to directly infer the desired result from [11], and we thus provide a detailed proof in Section 3, though we still naturally re-use the main tools and ideas proposed in [11].

Specialization to dense regimes. As hinted above, when the $\min\{T^*, n\}$ term is attained by n, our results indicate that one-by-one testing is asymptotically optimal. This is consistent with the above-mentioned result of [1], but also strengthens it in two ways:

- Individual testing is not only asymptotically optimal when the goal is to succeed with probability approaching one, but also when the goal is attaining *any* strictly positive target success probability.
- Individual testing is not only asymptotically optimal when $k = \Theta(n)$, but also when $k = \omega(\frac{n}{\log n})$, or even more generally, when $k > \frac{n\ln 2}{\log_2 n}$.

On the other hand, it is worth noting that our result only indicates failure when $T < (1-\epsilon)n$, whereas that of [1] handles the more general scenario $T < n-1$. This distinction is necessary when establishing high-probability failure and/or handling the regime $k = o(n)$, since otherwise one could consider a strategy that (e.g.) tests the first $n-2$ items one-by-one and then guesses the remaining two to be non-defective.

Note on partially concurrent work. In the initial version of our work, we focused only on the lower bound, and provided a weaker result with an unspecified coefficient to the $k\log n$ term in the $\min\{k\log n, n\}$ scaling. After releasing the initial version, the important case of $k = \Theta(\frac{n}{\log n})$ was studied in more detail

[2]The assumption $p = \omega(\frac{1}{n})$ ensures that $\overline{k} = np = \omega(1)$, and hence the number of defectives concentrates around $\overline{k}$. The theorem cannot be true as stated when $p \le O(\frac{1}{n})$; for example, even the trivial strategy of declaring every item non-defective has $\Omega(1)$ probability of succeeding, in contrast to the second part of the theorem. In addition, our main novel contribution is handling the significantly denser regime $\overline{k} = n^{1-o(1)}$.

in [17], giving upper and lower bounds with explicit constants. The updated version of our work was developed in parallel with [17], and establishes the precise constants. Our upper bound in fact matches that of [17] (and is proved similarly), whereas a refinement of the main proof technique is needed to obtain our tight lower bound (see the stopping condition of Step 4(a), Procedure 3.1).

2.2 Further Existing Results

Before proceeding, we provide a brief summary of some further existing works. Since these are less directly related to our work, we omit the details, and refer the reader to [5, 14] for more detailed surveys.

For certain variants of group testing, the optimal number of tests is $\Theta(k \log \frac{n}{k})$, as opposed to $\Theta(\min\{k \log n, n\})$ under the setup we consider. Specifically, two notable cases with scaling $\Theta(k \log \frac{n}{k})$ are (i) the adaptive setting, in which each test can be designed based on previous outcomes [2, 3, 20], and (ii) the approximate recovery criterion, in which $\Theta(k)$ false positives and $\Theta(k)$ false negatives are allowed in the reconstruction [27, 28].

In contrast to the noiseless setting that we consider in this paper, in the *noisy* setting, the number of tests is at least $\Omega(k \log n)$ even if $k \log n \gg n$, and even if adaptivity is allowed [26].

Finally, while the focus of our work is on high-probability recovery, extensive results have been established for the stronger guarantee of *uniform recovery*, i.e., a single test matrix that uniquely recovers any defective set of cardinality at most k, without allowing any error probability (e.g., see [9, 14, 15, 23] and the references therein). This stronger guarantee comes at the price of requiring significantly more tests, with a quadratic dependence on k instead of a linear dependence. In addition, the associated proof techniques are very different.

3 Proofs of Algorithm-Independent Lower Bounds

We first consider Theorem 2.1 regarding the i.i.d. prior, and then turn to Corollary 2.1 regarding the combinatorial prior.

3.1 Proof of the Lower Bound for Theorem 2.1

Our analysis builds on the ideas of [1, 11], both of which identify *totally disguised* items (see Definition 3.1 below) whose defectivity status can be flipped without changing the test outcomes. In [1], one such item suffices for attaining the weak converse (i.e., $\mathbb{P}[\hat{S} \neq S] \not\to 0$) in the linear regime. To obtain a stronger statement of the form $\mathbb{P}[\hat{S} \neq S] \to 1$ and also handle sublinear sparsity regimes, we follow the idea from [11] of identifying *many* such items.

Specifically, we follow the high-level steps of [11] and utilize certain auxiliary results therein, but modify the details in order to handle the regime $k = n^{1-o(1)}$ instead of $k \leq n^{1-\Omega(1)}$. The key idea is to identify many items that are disguised *independently of one another*. We then apply an auxiliary result of [1] (see Lemma 3.2 below) along with some "clean-up" steps to ensure that its assumptions remain valid each time it is invoked.

In the following, we let $q = 1 - p$ for convenience. The following useful definition was introduced in [1].

Definition 3.1. [1] We say that an item i is *disguised* in test t if at least one of the other items in the test is defective. We say that an item is *totally disguised* if it is disguised in every test it is included in. Let D_i denote the event that item i is totally disguised.

It is noted in [1] that if an item is totally disguised, then it remains totally disguised even if it is changed from defective to non-defective or vice versa. Thus, under the i.i.d. prior, the tests do not reveal any information about that item's defectivity status, and we have the following.

Lemma 3.1. (Implicit in [1] and [11, Sec. 3]) *For any given test matrix $\mathbf{X}$, and a defective set S generated according to the i.i.d. prior, we have the following: Conditioned on a given item i being totally disguised, that item is defective with conditional probability p (i.e., the same as the prior defectivity probability).*

It follows that for any totally disguised item, the best the algorithm can do is choose the more likely outcome, and succeed with probability $\max\{p, 1-p\} = 1-p$ (recalling that we focus on the case that $p \leq \frac{1}{2}$).

The following result from [1] is crucial for characterizing the probability of items being totally disguised.

Lemma 3.2. [1, Eq. (1)] *Define $\mathcal{L}(p) = \min_{x=2,3,\ldots,n} x\ln(1-q^{x-1})$, where $q = 1-p$. If the test design $\boldsymbol{X}$ has no tests with 0 or 1 items, then*

$$\frac{1}{n}\sum_{i=1}^{n} \ln \mathbb{P}[D_i] \geq \frac{T}{n} \cdot \mathcal{L}(p) \tag{3.1}$$

Hence, there exists an item i with $\ln \mathbb{P}[D_i] \geq \frac{T}{n} \cdot \mathcal{L}(p)$.

At a high level, this lemma is proved by directly calculating $\mathbb{P}[D_i]$ in terms of the i-th test size (which x plays the role of in the definition of $\mathcal{L}(p)$), then averaging over the resulting log-values and applying some simple lower bounding techniques.

In the linear regime (i.e., $k = \Theta(n)$), one has $\mathcal{L}(p) = \Theta(1)$, and consequently, Lemma 3.2 directly implies that the success probability is bounded away from one, after removing all tests with 0 or 1 items [1]. More generally, it is natural to ask whether there are, in fact, *many* items i with $\ln \mathbb{P}[D_i]$ close to the the right-hand side of (3.1) [11]. If we can find a "large" set W of such items such that these items are totally disguised independently from each other, then we may apply standard binomial distribution concentration bounds to conclude that many totally disguised items exist, with high probability.

Following [11], we interpret the testing strategy as a bipartite graph $G_{\mathbf{X}}$ in which there is a vertex v_i for each item i and a vertex v_t for each test t, with an edge between v_i and v_t if item i is placed in test t. Before constructing the desired set (denoted by W), we present two simple lemmas (which are analogous to [11, Lemmas 3.7 and 3.8]) and two subroutines that will be useful.

Lemma 3.3. *Let $z = \frac{2}{\ln \frac{1}{q}}$, and suppose that $T \leq n$. Then, the probability that there exists a negative test containing more than $z\ln n$ items is at most $\frac{1}{n}$.*

Proof. Recalling that $q = 1-p$, a given test containing at least $z\ln n$ items is negative with probability at most

$$q^{z\ln n} = e^{z\ln q \ln n} = \frac{1}{n^2} \tag{3.2}$$

by the definition of z. Since $T \leq n$, a union bound yields the desired result. □

We henceforth assume that no test contains more than $z\ln n$ items, since Lemma 3.3 implies that the decoder may declare all such tests to be positive without increasing the error probability by more than $\frac{1}{n} \to 0$.

Lemma 3.4. *Fix $\xi > 0$, and define an item to be* very-present *if it appears in more than n^{ξ} tests. If $T \leq n$ and no test contains more than $z\ln n$ items, then there are no more than $zn^{1-\xi}\ln n$ very-present items.*

Proof. We count the number P of pairs (i,t) such that item i is in test t. By assumption, $P \leq Tz\ln n \leq nz\ln n$. Letting n_{vp} be the number of very-present items, it follows that $n_{\mathrm{vp}}n^{\xi} \leq P < nz\ln n$, and rearranging yields the desired result. □

We are now in a position to describe the construction of the desired set W, namely, a set of items that are disguised independently of one another. To establish a hardness result, we would like to ensure that the size of W and the probability of each $i \in W$ being disguised are both large enough, so that the resulting error probability is high.

Towards achieving this goal, we introduce Subroutines 3.1 and 3.2. `Clean` removes all tests with 0 or 1 items, allowing us to apply Lemma 3.2, and `Extract` adds an item to W. Both will be called multiple times in the construction of W, and their calls will reduce the effective T and/or n.

The full procedure for constructing W is described in Procedure 3.1, which depends on a generic constant $\xi > 0$; although its use in step 4(a) is not directly related to its used in Lemma 3.4, we find it sufficient to

Subroutine 3.1: `Clean`($\mathbf{X}$).

1. Identify the set of tests $T_{\leq 1}$ containing 0 or 1 items, and the set of items I contained in at least one test in $T_{\leq 1}$.
2. Return $\mathbf{X}_{\geq 2}$, defined to be $\mathbf{X}$ with the rows and columns indexed by $T_{\leq 1}$ and I removed.

Subroutine 3.2: `Extract`($\mathbf{X}, W$).

1. Let $\widetilde{D}_i$ be the event that i is totally disguised with respect to $\mathbf{X}$. Let the item with the highest $\mathbb{P}[\widetilde{D}_i]$ be denoted by i_0, and set $W_{\text{next}} = W \cup \{i_0\}$.
2. Let T_{close} and I_{close} denote the sets of tests and items within distance at most 4 from i_0 in $G_{\mathbf{X}}$.
3. Set $\mathbf{X}_{\text{pruned}}$ to be $\mathbf{X}$ with the rows and columns indexed by T_{close} and I_{close} removed.
4. Return $(\mathbf{X}_{\text{pruned}}, W_{\text{next}})$

use the same constant in both cases. To justify step 1, we momentarily imagine that there exists a "genie" that tells the decoder the identity of the very-present items. Let the test results for G_0 and G_1 be $\mathbf{y}_0$ and $\mathbf{y}_1$ respectively; then, knowing $\mathbf{X}$, we see that $\mathbf{y}_0$ can be derived from $\mathbf{y}_1$ and the genie information. If we can prove that the error probability tends to one even with the help of the genie (and knowing $\mathbf{y}_1$), then it certainly tends to one without it, so step 1 is justified. After step 1, each item is contained in at most n^{ξ} tests.

Let w_i denote the i-th item placed in W. Let D_{w_i} be the event that w_i is totally disguised with respect to $\mathbf{X}_1$, and let $\widetilde{D}_{w_i}$ be the event that w_i is totally disguised with respect to $\mathbf{X}_{\text{tmp},i}$ (see Procedure 3.1 for the definitions of $\mathbf{X}_1$ and $\mathbf{X}_{\text{tmp},i}$). Since the totally disguised event D_{w_i} only depends on the 2-neighborhood of w_i in G_1, and the 2-neighborhoods of items in W are pairwise disjoint by construction (due to the `Extract` subroutine), the events $\{D_w : w \in W\}$ are independent (this independence property for nodes having distance greater than 4 was also used in [11]).

Next, we state the following simple lemma relating the events D_{w_i} and $\widetilde{D}_{w_i}$, both of which represent events of being totally disguised, but with respect to different test matrices.

Lemma 3.5. *Under the preceding setup, we have* $\mathbb{P}[D_{w_i}] \geq \mathbb{P}[\widetilde{D}_{w_i}]$.

Proof. In each `Clean`/`Extract` step, whenever we remove a test, we remove all of its items. It follows that w_i is contained in *the same tests* in $\mathbf{X}_1$ and $\mathbf{X}_{\text{tmp},i}$, except that each such test in $\mathbf{X}_{\text{tmp},i}$ has *fewer items*. Since a disguised item always remains disguised when further items are added to its tests, it follows that $\widetilde{D}_{w_i}$ implies D_{w_i}. □

In addition, we have the following lower bound on $|W|$, the total number of extracted items. Here and subsequently, we recall that to prove Theorem 2.1, it suffices to consider the regime $p = n^{-o(1)}$, since for any smaller p (i.e. $p = n^{-\Omega(1)}$), Theorem 2.1 was already established in [11].

Lemma 3.6. *Under the preceding setup, if* $p = n^{-o(1)}$ *and* $T \leq (1-\epsilon)n$*, then the size of the set* W *returned by Procedure 3.1 satisfies the following:*

$$|W| \geq n^{1-3\xi}. \tag{3.3}$$

Proof. We first count the number of removed items as follows:

- No more than T items alone in some test are removed by `Clean`.
- Lemma 3.4 implies that we removed at most $zn^{1-\xi}\ln n$ very-present items, and this scales as $o(n)$ due to the fact that $z = \frac{2}{\ln \frac{1}{1-p}} = \Theta(\frac{1}{p}) = n^{o(1)}$ (by the assumption $p = n^{-o(1)}$).

Procedure 3.1: `ConstructSet`($\mathbf{X}$).

1. Let $G_0 = G_{\mathbf{X}}$, and let (n, T) be the number of items and tests in $\mathbf{X}$. Remove all very-present items from G_0 to obtain G_1. Let $G = G_1$.
2. Initialize $W_0 = \emptyset$, $i = 1$.
3. Set $\mathbf{X}_i \leftarrow$ test design represented by G_i. Set $\mathbf{X}_{\mathrm{tmp},i} \leftarrow \texttt{Clean}(\mathbf{X}_i)$, and let (n_i, T_i) be the corresponding number of items and tests in $\mathbf{X}_{\mathrm{tmp},i}$.
4. Perform the following:
   (a) If $n_i > 0$ and $\frac{T_i}{n_i} \le (1+\xi)\frac{T}{n}$, then set $(\mathbf{X}_{i+1}, W_i) \leftarrow \texttt{Extract}(\mathbf{X}_{\mathrm{tmp},i}, W_{i-1})$, $G_{i+1} \leftarrow G_{\mathbf{X}_{i+1}}$, and $i \leftarrow i+1$, and return to Step 3.
   (b) Otherwise, terminate the procedure and return $W = W_{i-1}$.

- By the assumption stated following Lemma 3.3 and the removal of very-present items, each call to `Extract` removes at most $z^2 n^{2\xi} \ln^2 n$ items.

We now argue by contradiction that (3.3) must hold. Suppose to the contrary that Procedure 3.1 terminates at some iteration $i^* \le n^{1-3\xi}$. Then, the above calculations imply that

$$n_{i^*} \ge n - T - zn^{1-\xi} \ln n - n^{1-3\xi} \cdot z^2 n^{2\xi} \ln^2 n \tag{3.4}$$

$$\ge \epsilon n - o(n), \tag{3.5}$$

where we used the fact that $T \le (1-\epsilon)n$ and $z = n^{o(1)}$. This means that the stopping condition met in step 4(a) cannot have been n_i reaching zero, so it must have been $\frac{T_{i^*}}{n_{i^*}}$ exceeding $(1+\xi)\frac{T}{n}$.

While (3.5) indicates that the majority of items could eventually be removed in principle, this is only due to the subtraction of T in (3.4); the other two terms behave as $o(n)$, and we conclude that the removal of very-present items and the calls to `Extract` collectively only remove $o(n)$ items. Any further removal of items can only be due to the tests containing one item in `Clean`; removing these causes T_i and n_i to be reduced by the *same amount*. However, as long as $\frac{T_i}{n_i} < 1$ (which holds by assumption for $i = 0$, and subsequently for all $i \le i^*$ due to the stopping condition), reducing T_i and n_i by the same amount can *only make the ratio smaller* (i.e., $\frac{T_i - c}{n_i - c} \le \frac{T_i}{n_i}$ for any $c \in [0, T_i]$).

More formally, suppose that up to index i^*, a total of c items and tests are removed due to tests containing a single item, and a total of c' items are removed for the other reasons mentioned above. Then, we have

$$\frac{T_{i^*}}{n_{i^*}} \le \frac{T - c}{n - c - c'}. \tag{3.6}$$

As established above, we have $0 \le c \le T \le n(1-\epsilon)$ and $c' = o(n)$. However, since $\frac{T}{n} < 1$ by assumption, substituting these findings into (3.6) reveals that $\frac{T_{i^*}}{n_{i^*}} \le \frac{T}{n}(1 + o(1))$, which gives the desired contradiction to the stopping condition $\frac{T_{i^*}}{n_{i^*}} > (1+\xi)\frac{T}{n}$.

□

Recall that all of the extracted items have independent totally disguised events, each with probability lower bounded according to Lemma 3.2. We need to consider applying this lemma with possibly smaller choices of T and n than the original values (namely, T_i and n_i), but the stopping condition in step 4(a) of Procedure 3.1 ensures that $\frac{T_i}{n_i} \le (1+\xi)\frac{T}{n}$. As a result, Lemmas 3.2 and 3.5 guarantee for any extracted item i that

$$\mathbb{P}[D_i] \ge \exp\left(\frac{T(1+\xi)}{n} \cdot \mathcal{L}(p)\right), \tag{3.7}$$

where we recall that $\mathcal{L}(p) = \min_{x=2,3,\ldots,n} x \ln(1 - q^{x-1})$ with $q = 1 - p$. Note that $x \ln(1 - q^{x-1}) < 0$, so this minimum is to be interpreted as "most negative". This minimum is characterized in the following lemma, which is similar to [11, Claim 3.12].

Lemma 3.7. *For any $p \le \frac{1}{2}$ satisfying $p = n^{-o(1)}$, we have the following: (i) If $p = o(1)$, then $-\mathcal{L}(p) = \frac{(\ln 2)^2}{p}(1 + o(1))$; (ii) If $p = \Theta(1)$, then $-\mathcal{L}(p) = \Theta(1)$.*

Proof. We provide a simple generalization of the argument from [11, Claim 3.12], which focuses on the regime $k \le n^{1-\Omega(1)}$. We first write

$$-\mathcal{L}(p) = \max_{x=2,3,\dots,n} x \ln \frac{1}{1-q^{x-1}}. \tag{3.8}$$

This quantity is lower bounded by the argument corresponding to $x = \lceil \frac{1}{p} \rceil + 1$, which readily yields $\ln \frac{1}{1-q^{x-1}} = \Theta(1)$ and hence an $\Omega(\frac{1}{p})$ lower bound on $-\mathcal{L}(p)$.

For the upper bound, we note that for $x = o(\frac{1}{p})$ we have $q^{x-1} = (1-p)^{x-1} = 1 - \Theta(px)$, so the objective function behaves as $O(x \ln \frac{1}{px})$, which is $o(\frac{1}{p})$ (since $px \ln \frac{1}{px} \to 0$ as $px \to 0$). On the other hand, if $x = \omega(\frac{1}{p})$, then $q^{x-1} = (1-p)^{x-1} \to 0$, so the objective behaves as $O(x(1-p)^{x-1}) = O(xe^{-px})$, which is $o(\frac{1}{p})$ (since $pxe^{-\Theta(px)} \to 0$ as $px \to \infty$). Hence, the optimal choice of x must scale as $\Theta(\frac{1}{p})$, and in this case, we have $\ln \frac{1}{1-q^{x-1}} = \Theta(1)$, yielding an $O(\frac{1}{p})$ upper bound on $-\mathcal{L}(p)$.

The second part of the lemma follows immediately, whereas for the first part, a refined analysis is needed. For $p = o(1)$ and $x = \Theta(\frac{1}{p})$, we have

$$x \ln \frac{1}{1-(1-p)^{x-1}} = x \ln \frac{1}{1 - e^{-px}(1 + O(p))} \tag{3.9}$$

$$= -x \ln(1 - e^{-px}) + O(px) \tag{3.10}$$

by standard Taylor expansions. Since $O(px) = O(1)$ is asymptotically negligible compared to the $\Theta(\frac{1}{p})$ scaling derived above, it suffices to consider maximizing the first term. As noted in [11], we can define $d = px$ and write this term as $\frac{1}{p}(-d \ln(1 - e^{-d}))$, and it is a simple differentiation exercise to verify that $-d \ln(1 - e^{-d})$ is maximized at $d = \ln 2$, with maximum value $(\ln 2)^2$. While the corresponding choice $x = \frac{d}{p}$ may not be integer-valued, the effect of rounding is asymptotically negligible for $p = o(1)$ by the continuity of the function $-d \ln(1 - e^{-d})$. □

Combining Lemma 3.7 with (3.7), we obtain for some $c_p = \Theta(1)$ that

$$\mathbb{P}[D_i] \ge \exp\left(-\frac{(1+\xi)c_p T}{np}\right), \tag{3.11}$$

and moreover, when $p = o(1)$ we specifically have $c_p = (\ln 2)^2(1 + o(1))$.

Since the events $\{D_i\}_{i \in W}$ are mutually independent by construction, and $|W| \ge n^{1-3\xi}$ according to Lemma 3.6, we deduce that the number of totally disguised items is stochastically dominated by Binomial$(n^{1-3\xi}, e^{-\frac{(1+\xi)c_p T}{np}})$. In particular, the average number of totally disguised items is

$$n^{1-3\xi} e^{-\frac{(1+\xi)c_p T}{np}}, \tag{3.12}$$

and by simple re-arrangements, this is lower bounded by n^{ξ} whenever

$$T \le \frac{np(1-4\xi)}{(1+\xi)c_p} \ln n. \tag{3.13}$$

Thus, by the multiplicative form of the Chernoff bound, the actual number is at least $N_{\min} := \frac{1}{2}n^{\xi}$ with probability approaching one when (3.13) holds.

By Lemma 3.1 and the assumption $p \le \frac{1}{2}$, for any item that is disguised, the optimal algorithm can do no better than declare it to be non-defective, and the resulting probability of being correct is at most

$$(1-p)^{N_{\min}} \le e^{-pN_{\min}} = e^{-\frac{p}{2}n^{\xi}} = o(1), \tag{3.14}$$

where the last step follows from the assumption $p = n^{-o(1)}$. Thus, we have proved that $\mathbb{P}[\hat{S} = S] = o(1)$ whenever $T \le (1-\epsilon)n$ and (3.13) holds.

When $p = \Theta(1)$ (or more generally $p = \omega(\frac{1}{\log n})$), the stricter of these two conditions is $T \leq (1-\epsilon)n$, and the constant c_p in (3.13) is inconsequential. On the other hand, when $p = o(1)$, we have established that $c_p = (\ln 2)^2(1+o(1))$. Since ξ can be arbitrarily small, it follows that (3.13) reduces to $T \leq (1-\epsilon')\frac{np}{(\ln 2)^2}\ln n$ for arbitrarily small $\epsilon' > 0$. Finally, since $\frac{\ln n}{\ln 2} = \log_2 n$, and the assumption $p = n^{-o(1)}$ implies that $\log_2 n = (\log_2 k)(1+o(1))$, we obtain the desired threshold corresponding to the second term of T^* in (2.1).

3.2 Proof of the Lower Bound for Corollary 2.1

We utilize an approach from [11, Lemma 3.6] for transferring the key auxiliary results on the number of disguised items from the i.i.d. prior to the combinatorial prior. Despite the high level of similarity, we provide the main details for completeness.

The idea is to show that with too few tests, the number of totally disguised defectives and totally disguised non-defectives both grow unbounded with high probability. When this occurs, interchanging the statuses among these items would not impact the test results, and hence, there exist an unbounded number of candidate defective sets of cardinality k consistent with the test outcomes. The decoder cannot do any better than guess one of these at random, failing with high probability. This intuition is easily made precise [11], giving the following.

Lemma 3.8. [11, Facts 3.1 and 3.3] *Under the combinatorial prior, the conditional error probability of any group testing strategy given that there are $\tilde{n}_0$ totally disguised non-defectives and $\tilde{n}_1$ totally disguised defectives is at least $1-\frac{1}{\tilde{n}_0\tilde{n}_1}$. In particular, if $\tilde{n}_0 = \omega(1)$ and $\tilde{n}_1 = \omega(1)$, then the conditional error probability is $1-o(1)$.*

Consider the combinatorial prior with $n^{1-o(1)} \leq k \leq \frac{n}{2}$, where the condition $k \geq n^{1-o(1)}$ is safe to assume since Corollary 2.1 is already well-known when $k \leq n^{1-\Omega(1)}$ (see Section 2). We consider generating S according to the following procedure:

1. Let $S_0 \subseteq [n]$ include each item independently with probability $p_0 = \frac{k-\sqrt{k}\ln n}{n}$. That is, S_0 follows the i.i.d. prior with parameter p_0.
2. Form S by adding $\max\{k-|S_0|, 0\}$ elements of $[n] \setminus S_0$ to S_0, chosen uniformly at random.

By the symmetry of this construction, conditioned on the event $|S_0| \leq k$, the resulting set S is indeed distributed according to the combinatorial prior. While $|S_0| > k$ has a non-zero probability, for the purposes of proving a converse, we can simply assume that this event always leads to successful recovery. Since we assume that $k \geq n^{1-o(1)}$, a simple concentration argument (e.g., the Chernoff bound or central limit theorem) gives with probability $1-o(1)$ that

$$k - 2\sqrt{k}\ln n \leq |S_0| \leq k, \tag{3.15}$$

so the resulting contribution to the success probability is asymptotically negligible.

We now introduce the terminology that an item i is *totally disguised in the first step* if the defectives from S_0 alone are enough to disguise i in every test it is included in. Clearly, being totally disguised in the first step is sufficient for being totally disguised after the second step, since the second step only involves marking more items as defective.

Hence, trivially, the number of totally disguised defective items only increases (or stays the same) after the second step. The number of totally disguised non-defectives may in principle decrease due to non-defectives being changed to defective, but conditioned on (3.15), any given non-defective is only changed with probability $O(\frac{\sqrt{k}\ln n}{n}) = o(1)$. As a result, if there are $\omega(1)$ totally disguised non-defectives, the same still remains true with probability $1-o(1)$ after the second step.

Hence, in accordance with Lemma 3.8, it suffices to show that under the i.i.d. prior with parameter $p_0 = \frac{k-\sqrt{k}\ln n}{n}$, the number of totally disguised defectives and totally disguised non-defectives both behave as $\omega(1)$ with probability $1-o(1)$. Note that the assumption $n^{1-o(1)} \leq k \leq \frac{n}{2}$ ensures that $n^{-o(1)} \leq p_0 \leq \frac{1}{2}$, as was assumed in the later parts of Section 3.1.

We already argued that when (3.13) holds, the average number of totally disguised items is at least n^{ξ}. Since $n^{-o(1)} \leq p_0 \leq \frac{1}{2}$, it follows that the average number of totally disguised defectives and totally

disguised non-defectives are both at least $n^{\xi-o(1)}$ on average. Again, the multiplicative form of the Chernoff bound implies the same with high probability, and we have the desired $\omega(1)$ scaling. This establishes that the condition on T from Theorem 2.1 with p_0 in place of p is necessary for attaining a success probability bounded away from zero, and since $np_0 = k(1+o(1))$ by definition, Corollary 2.1 follows.

4 Conclusion

We have proved that the optimal number of tests for probabilistic noiseless non-adaptive group testing is $\Theta(\min\{k \ln n, n\})$, as well as establishing the precise underlying constant factors. This closes gaps exhibited by existing bounds in the in the case that k is "mildly" sublinear in n, so that the optimal thresholds are now known for arbitrary scaling regimes. Perhaps the main challenge remaining in this setting is to understand how the number of tests increases when the target error probability decreases to zero at a given rate depending on n.

A Proofs of Algorithmic Upper Bounds

Under the combinatorial prior, the algorithmic upper bound for Theorem 2.1 in the regime $k \leq n^{1-\Omega(1)}$ is proved in [22] using the following strategy:

- **Test design.** Generate the $T \times n$ test matrix $\mathbf{X}$ according to the *near-constant tests-per item* design: For each item $i = 1, \dots, n$, select $L = \lfloor \frac{T \ln 2}{k} \rfloor$ tests uniformly at random *with replacement*, and set the corresponding entries in the i-th column of $\mathbf{X}$ to one.
- **Decoding algorithm.** Given the test outcomes, estimate the defective set using the *Definite Defectives (DD)* algorithm:
  - (i) Mark all items in negative tests as *definitely non-defective*, and all remaining items as *possibly defective* (PD);
  - (ii) For any PD item appearing in some (necessarily positive) test without any other PD items, mark it as *definitely defective* (DD).
  - (iii) Return the set of DD items as the final estimate.

The main result of [22] states that when $k = \Theta(n^\theta)$ with $\theta \in (0,1)$ and the number of tests satisfies

$$T \geq \frac{\max\{\theta, 1-\theta\}}{\ln 2}(k \log_2 n)(1+\epsilon) \tag{A.1}$$

for some $\epsilon > 0$, the resulting error probability approaches zero as $n \to \infty$. Our goal is to generalize this result to denser sparsity regimes.

The condition (A.1) ensures that $L = \lfloor \frac{T \ln 2}{k} \rfloor$ scales as $\omega(1)$. Hence, the effect of rounding is negligible, in the sense that $L = \frac{T \ln 2}{k}(1+o(1))$. As in [22], we subsequently work with the exact expression $L = \frac{T \ln 2}{k}$ for notational convenience, since the $o(1)$ term does not affect the final result.

With the regime $k \leq n^{1-\Omega(1)}$ having been handled in [22], it suffices to consider $k = n^{1-o(1)}$. In this regime, it holds that $T^*(n,k) = \frac{k \log_2 k}{\ln 2}$. For convenience, we apply the fact that $\log_2 k = (\log_2 n)(1+o(1))$ (whenever $k = n^{1-o(1)}$), meaning that it suffices to show that the success probability approaches one when

$$T \geq \frac{k \log_2 n}{\ln 2}(1+\epsilon). \tag{A.2}$$

Observe that this matches (A.1), but with the quantity

$$m = \max\{\theta, 1-\theta\} \tag{A.3}$$

replaced by $m = 1$.

Analysis. We start with the following bound which is central to the analysis of [22], and is conveniently non-asymptotic so can can also be used here: For any defective $i \in S$, denoting the final estimate by $\widehat{S}$, we have

$$\mathbb{P}[i \notin \widehat{S}] \leq \underbrace{\sum_{w \in [w_-, w_+]} \mathbb{P}[W^{(S\setminus i)} = w] \sum_{j=0}^{L} \mathbb{P}[M_i = j | W^{(S\setminus i)} = w] \phi_j(1/w_-, g^* L)}_{:=\Psi_1} + \underbrace{\mathbb{P}[W^{(S\setminus i)} \notin [w_-, w_+]]}_{:=\Psi_2} + \underbrace{\mathbb{P}[G > g^* | W^{(S\setminus i)} \notin [w_-, w_+]]}_{:=\Psi_3}, \tag{A.4}$$

where:

- $W^{(S\setminus i)}$ denotes the number of (necessarily positive) tests containing at least one item in $S \setminus \{i\}$, i.e., a defective item differing from i;
- M_i denotes the number of tests containing $i \in S$ and no other defectives;
- G denotes the number of non-defectives that do not appear in any negative tests;
- w_- and w_+ are arbitrary thresholds, but should be chosen to ensure that $W^{(S\setminus i)} \in [w_-, w_+]$ with high probability;
- g^* is an arbitrary threshold, but should be chosen to ensure that $G \leq g^*$ with high probability;
- $\phi_j(s, V) = \sum_{\ell=0}^{j} (-1)^\ell \binom{j}{\ell} (1 - \ell s)^V$ is a quantity arising from applying the inclusion-exclusion principle to a union of events in a coupon collector problem [22, Appendix B].

We set w_-, w_+, and g^* in the same way as [22]:

$$w_- = \frac{T}{2}(1 - \delta) \tag{A.5}$$

$$w_+ = \frac{T}{2}(1 + \delta) \tag{A.6}$$

$$g^* = n\Big(\frac{1}{2} + \delta\Big)^L, \tag{A.7}$$

for some $\delta > 0$ to be specified later. The interaction between δ and ϵ (see (A.2)) turns out to be slightly delicate, and choosing them appropriately is the main difference here compared to [22].

The analysis of [22] focuses on the case that (A.1) holds with equality. This is without loss of generality, since additional tests can only ever help the DD algorithm. We similarly assume that (A.2) holds with equality. In view of the union bound over the k defectives, the goal is to show that $k\Psi_\nu \to 0$ for $\nu \in \{1, 2, 3\}$ in (A.4). We proceed as follows:

1. For Ψ_1, it is shown in [22, Eq. (39)] that if $L = m(1 + \epsilon)\frac{\ln n}{\ln 2}$ (which holds via $L = \frac{T \ln 2}{k}$ and equality holding in (A.1)) and $k \leq cn^m$ for some constant c (with m given in (A.3)), then

$$k\Psi_1 \leq c \exp\Big(\frac{L^2}{4w_-}\Big) \exp\bigg(-\Big(\epsilon - \frac{1 + \epsilon}{\ln 2}\Big(\delta + \frac{g^* L}{w_-}(1 - \delta)\Big)\Big) m \ln n\bigg). \tag{A.8}$$

   Recall that we are adopting the choice $m = 1$; this means that the condition $k \leq cn^m$ is trivially satisfied with $c = 1$. Hence, if we can further establish that $\frac{L^2}{4w_-} = o(1)$ and $\frac{g^* L}{w_-} = o(1)$, it will follow from (A.8) that

$$k\Psi_1 \leq (1 + o(1)) \exp\bigg(-\Big(\epsilon - \frac{1 + \epsilon}{\ln 2}(\delta + o(1))\Big) \ln n\bigg). \tag{A.9}$$

   This approaches zero as $n \to \infty$ when δ is strictly smaller than $\frac{\epsilon \ln 2}{1+\epsilon}$. For concreteness, we set $\delta = \frac{2}{3}\epsilon$ (note that $\frac{2}{3} < \ln 2$), so that the preceding requirement holds when ϵ is sufficiently small.

The above-mentioned requirement $\frac{L^2}{4w_-} = o(1)$ follows immediately from the fact that $L = \Theta(\log n)$ and $w_- = \Theta(T) = \Theta(k \log n)$ (with $k = n^{1-o(1)}$). As for $\frac{g^* L}{w_-}$, the steps in [22, Eq. (31)] turn out to be too loose for our purposes, but are easily modified: Combining $L = \frac{T \ln 2}{k}$ with (A.5) gives $\frac{L}{w_-} = \frac{2 \ln 2}{k(1-\delta)}$, and further combining with (A.7) gives

$$\ln \frac{g^* L}{w_-} = \ln \frac{n}{k} + L \ln \Big(\frac{1}{2} + \delta \Big) + \ln \frac{2 \ln 2}{1 - \delta}. \tag{A.10}$$

The assumption $k = n^{1-o(1)}$ gives $\ln \frac{n}{k} = o(\log n)$, and combining this with $L = \Theta(\log n)$ and $\ln(\frac{1}{2} + \delta) < 0$ (for small enough δ), it follows that the right-hand side of (A.10) approaches $-\infty$, and hence $\frac{g^* L}{w_-} = o(1)$ as desired.

2. For Ψ_2, we can directly use the following finding from [22] based on McDiarmid's inequality:

$$k\Psi_2 \le k \exp \Big(\frac{\delta^2 T}{4 \ln 2} (1 + o(1)) \Big). \tag{A.11}$$

This approaches zero as $n \to \infty$, since $T = \Theta(k \log n)$.

3. For Ψ_3, we use the following bound [22, Eq. (42)] based on Bernstein's inequality, which holds provided that $L \to \infty$ (which we already established) and $\delta \le \frac{1}{4}$:

$$k\Psi_3 \le k \exp \bigg(- n \frac{(1/2 + \delta)^L}{2/3 + o(1)} \bigg). \tag{A.12}$$

Recall that $L = \frac{T \ln 2}{k}$; substituting T equaling the right-hand side of (A.2) gives $L = (1 + \epsilon) \log_2 n$. We proceed by considering the logarithm (base 2) of $n(1/2 + \delta)^L$:

$$\log_2 \Big(n(1/2 + \delta)^L \Big) = \log_2 n + L \log_2 \Big(\frac{1}{2} + \delta \Big) \tag{A.13}$$

$$= (\log_2 n) \bigg[1 + (1 + \epsilon) \log_2 \Big(\frac{1}{2} + \delta \Big) \bigg]. \tag{A.14}$$

Using the above choice $\delta = \frac{2}{3}\epsilon$, a simple Taylor expansion yields the following as $\epsilon \to 0$:[3]

$$(1 + \epsilon) \log_2 \Big(\frac{1}{2} + \frac{2}{3}\epsilon \Big) = -1 + \epsilon \Big(\frac{2}{3} \cdot \frac{2}{\ln 2} - 1 \Big) + o(\epsilon). \tag{A.15}$$

Hence, since $\frac{2}{3} \cdot \frac{2}{\ln 2} \approx 1.92 > 1$, we have for sufficiently small ϵ that (A.14) is positive and scales as $\Theta(\log n)$, and substituting into (A.12) gives $k\Psi_3 \le k \exp(-n^{\Theta(1)}) \to 0$, as desired.

Since the above analysis holds for arbitrarily small $\epsilon > 0$ (and hence arbitrarily small $\delta > 0$ via $\delta = \frac{2}{3}\epsilon$) when the number of tests satisfies (A.2) with equality, the upper bound in Theorem 2.1 follows.

Handling the i.i.d. prior. While [22] only considers the combinatorial prior with a fixed value of k, the analogous result follows essentially immediately for the i.i.d. prior, in which k is a random variable. This is because by a simple concentration argument (e.g., Hoeffding's inequality), as long as $np \to \infty$, it holds that $k = np(1 + o(1))$ with probability approaching one. We can therefore replace the choice $L = \frac{T \ln 2}{k}(1 + o(1))$ by $L = \frac{T \ln 2}{np}(1 + o(1))$, and under the high-probability event $k = np(1 + o(1))$, the two are equivalent up to a change in the $o(1)$ term. Since conditioning on any particular value of k under the i.i.d. prior brings us back to the combinatorial prior, the desired result follows.

Acknowledgment

We are very grateful to an anonymous reviewer for helpful suggestions and encouraging us to establish the precise constant factors in the analysis, which allowed us to significantly strengthen an earlier version of our main result containing an unspecified constant.

[3] In fact, a visual plot reveals that $1 + (1 + \epsilon) \log_2 (\frac{1}{2} + \frac{2}{3}\epsilon)$ is positive for all $\epsilon > 0$.

References

[1] Aldridge, M. (2019) Individual Testing Is Optimal for Nonadaptive Group Testing in the Linear Regime. *IEEE Trans. Inf. Theory*, **65**(4), 2058–2061.

[2] Aldridge, M. (2019) Rates of adaptive group testing in the linear regime. in *IEEE Int. Symp. Inf. Theory (ISIT)*.

[3] ——— (2020) Conservative two-stage group testing. https://arxiv.org/abs/2005.06617.

[4] Aldridge, M., Baldassini, L. & Johnson, O. (2014) Group Testing Algorithms: Bounds and Simulations. *IEEE Trans. Inf. Theory*, **60**(6), 3671–3687.

[5] Aldridge, M., Johnson, O. & Scarlett, J. (2019) Group Testing: An Information Theory Perspective. *Found. Trend. Comms. Inf. Theory*, **15**(3–4), 196–392.

[6] Atia, G. & Saligrama, V. (2012) Boolean Compressed Sensing and Noisy Group Testing. *IEEE Trans. Inf. Theory*, **58**(3), 1880–1901.

[7] Baldassini, L., Johnson, O. & Aldridge, M. (2013) The capacity of adaptive group testing. in *IEEE Int. Symp. Inf. Theory*, pp. 2676–2680.

[8] Chan, C. L., Che, P. H., Jaggi, S. & Saligrama, V. (2011) Non-adaptive probabilistic group testing with noisy measurements: Near-optimal bounds with efficient algorithms. in *Allerton Conf. Comm., Ctrl., Comp.*, pp. 1832–1839.

[9] Cheraghchi, M. (2013) Noise-resilient group testing: Limitations and constructions. *Disc. App. Math.*, **161**(1), 81–95.

[10] Coja-Oghlan, A., Gebhard, O., Hahn-Klimroth, M. & Loick, P. (2019) Information-theoretic and algorithmic thresholds for group testing. in *Int. Colloq. Aut., Lang. and Prog. (ICALP)*.

[11] ——— (2020) Optimal group testing. in *Conf. Learn. Theory (COLT)*.

[12] Cormode, G. & Muthukrishnan, S. (2005) What's Hot and What's Not: Tracking Most Frequent Items Dynamically. *ACM Trans. Database Sys.*, **30**(1), 249–278.

[13] Dorfman, R. (1943) The detection of defective members of large populations. *Ann. Math. Stats.*, **14**(4), 436–440.

[14] Du, D. & Hwang, F. K. (2000) *Combinatorial group testing and its applications*, vol. 12. World Scientific.

[15] D'yachkov, A. G. & Rykov, V. V. (1982) Bounds on the length of disjunctive codes. *Problemy Peredachi Informatsii*, **18**(3), 7–13.

[16] Fernández Anta, A., Mosteiro, M. A. & Ramón Muñoz, J. (2011) Unbounded Contention Resolution in Multiple-Access Channels. in *Distributed Computing*, vol. 6950, pp. 225–236. Springer Berlin Heidelberg.

[17] Flodin, L. & Mazumdar, A. (2021) Probabilistic Group Testing with a Linear Number of Tests. https://arxiv.org/abs/2106.06878 (extended version of ISIT 2021 paper).

[18] Freidlina, V. L. (1975) On a design problem for screening experiments. *Theory of Prob. & Apps.*, **20**(1), 102–115.

[19] Gilbert, A., Iwen, M. & Strauss, M. (2008) Group testing and sparse signal recovery. in *Asilomar Conf. Sig., Sys. and Comp.*, pp. 1059–1063.

[20] Hwang, F. (1972) A method for detecting all defective members in a population by group testing. *J. Amer. Stats. Assoc.*, **67**(339), 605–608.

[21] Inan, H. A., Kairouz, P., Wootters, M. & Özgür, A. (2019) On the Optimality of the Kautz-Singleton Construction in Probabilistic Group Testing. *IEEE Trans. Inf. Theory*, **65**(9), 5592–5603.

[22] Johnson, O., Aldridge, M. & Scarlett, J. (2019) Performance of group testing algorithms with near-constant tests-per-item. *IEEE Trans. Inf. Theory*, **65**(2), 707–723.

[23] Kautz, W. & Singleton, R. (1964) Nonrandom binary superimposed codes. *IEEE Trans. Inf. Theory*, **10**(4), 363–377.

[24] Malyutov, M. (1978) The separating property of random matrices. *Math. Notes Acad. Sci. USSR*, **23**(1), 84–91.

[25] Riccio, L., Colbourn, C. J. et al. (2000) Sharper bounds in adaptive group testing. *Taiwanese Journal of Mathematics*, **4**(4), 669–673.

[26] Scarlett, J. (2019) Noisy Adaptive Group Testing: Bounds and Algorithms. *IEEE Trans. Inf. Theory*, **65**(6), 3646–3661.

[27] Scarlett, J. & Cevher, V. (2016) Phase transitions in group testing. in *Proc. ACM-SIAM Symp. Disc. Alg. (SODA)*.

[28] ——— (2017) How little does non-exact recovery help in group testing?. in *IEEE Int. Conf. Acoust. Sp. Sig. Proc. (ICASSP)*.

[29] Ungar, P. (1960) The cutoff point for group testing. *Communications on Pure and Applied Mathematics*, **13**(1), 49–54.

[30] Yelin, I., Aharony, N., Tamar, E. S., Argoetti, A., Messer, E., Berenbaum, D., Shafran, E., Kuzli, A., Gandali, N., Shkedi, O. et al. (2020) Evaluation of COVID-19 RT-qPCR test in multi sample pools. *Clinical Infectious Diseases*, **71**(16), 2073–2078.