Noisy Group Testing in the Linear Regime Exact Thresholds and Efficient Algorithms

Lukas Hintze LUKAS.RASMUS.HINTZE@UNI-HAMBURG.DE
University of Hamburg, Department of Informatics, Vogt-Kölln-Str. 30, 22527 Hamburg, Germany

Lena Krieg LENA.KRIEG@TU-DORTMUND.DE
TU Dortmund, Faculty of Computer Science, Otto-Hahn-Str. 12, 44227 Dortmund, Germany

Olga Scheftelowitsch OLGA.SCHEFTELOWITSCH@TU-DORTMUND.DE
TU Dortmund, Faculty of Computer Science, Otto-Hahn-Str. 12, 44227 Dortmund, Germany

Haodong Zhu H.ZHU1@TUE.NL
Eindhoven University of Technology, Department of Mathematics and Computer Science, P.O. Box 513, Eindhoven, 5600 MB, The Netherlands

Editors: Nika Haghtalab and Ankur Moitra

Abstract

In group testing, the task is to identify defective items by testing groups of them together using as few tests as possible. We consider the setting where each item is defective with a constant probability α, independent of all other items. In the (over-)idealized noiseless setting, tests are positive exactly if any of the tested items are defective. We study a more realistic model in which observed test results are subject to noise, i.e., tests can display false positive or false negative results with constant positive probabilities. We determine precise constants c such that $cn \ln n$ tests are required to recover the infection status of every individual for both adaptive and non-adaptive group testing: in the former, the selection of groups to test can depend on previously observed test results, whereas it cannot in the latter. Additionally, for both settings, we provide efficient algorithms that identify all defective items with the optimal amount of tests with high probability. Thus, we completely solve the problem of binary noisy group testing in the studied setting.

The version containing all proofs can be found on arXiv (Hintze et al. (2024)).

1. Introduction

Originally motivated by large-scale screening of soldiers for syphilis in the Second World War, the study of group testing dates back to the 1940s Dorfman (1943). But its relevance exceeds its original inspiration by far, emphasized by the vast diversity of applications of group testing to other areas. Aldridge et al. (2019); Cheraghchi and Nakos (2020) list many applications, among them multiple access communication Wolf (1985), data compression and storage Hong et al. (2001), secure key distribution Chen et al. (2007), and quality control Sobel and Groll (1959). Even though this problem has been of interest for over 80 years, and despite being easy to state, it is far from being completely understood, adding to the appeal of group testing.

To describe the general problem setting, consider a large population of individuals, some of which are infected. The aim is to identify infected individuals such that the number of tests is minimized. To achieve this, one pools groups of individuals together, testing a

whole group with just one test. In an idealized *noiseless* setting, a test renders a positive result if and only if at least one participating individual is infected. However, in reality, tests can show false positive and false negative results Haar et al. (2011); Toptan et al. (2021); Jasima et al. (2022); Kılbaş et al. (2022). We study a *noisy* setting where tests err independently, i.e., a should-be-positive test is observed as negative with probability p_{10}, and a should-be-negative result is observed as positive with probability p_{01}. In other words, idealized test results are distorted by a noisy binary channel. In all group testing settings the main point of interest is, how many tests precisely are required to achieve a correct diagnosis of the individuals.

A significant proportion of recent literature, discussed in detail below, studies the *sparse regime*, where the number of infected people is *sublinear* in the population size. Noteworthy recent achievements include the complete understanding of noiseless group testing Coja-Oghlan et al. (2021) as well as the almost complete understanding of noisy group testing Coja-Oghlan et al. (2024); Chen and Scarlett (2024), both in the sparse regime. In contrast, results for the linear regime studied here lag far behind. For the noisy setting, they boil down to a non-tight lower bound on the number of necessary tests for the special case of symmetric noise Scarlett (2019). The lack of rigorous results in the linear regime is in stark contrast to the substantial amount of potential applications. E.g., considering medical conditions, a constant prevalence is most intuitive (the early stages of a pandemic being a notable exception). Moreover, the actually available tests are far from being perfectly accurate Jasima et al. (2022); Kılbaş et al. (2022). Of course, interest in noisy settings extends to the aforementioned variety of applications of group testing. Considering this, one might call noisy linear group testing one of the most realistic scenarios studied so far. Quite surprisingly, the question of the number of tests required precisely in this most realistic scenario has long been unanswered, despite group testing being an active area of research.

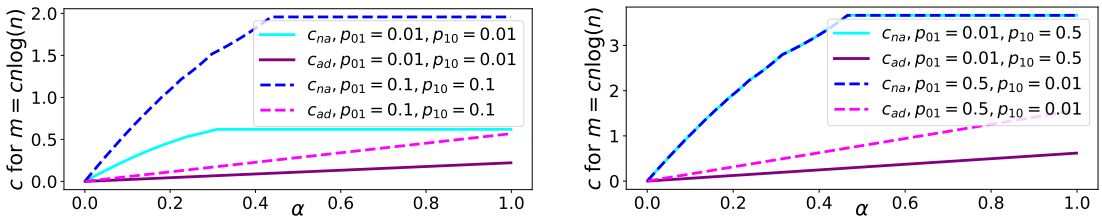
In group testing, one differentiates between adaptive and non-adaptive strategies: In the former, tests are conducted in (multiple) rounds, where the choice of tests in one round may depend on the results of previous tests. For the latter, there is only one round, so that the test design is fixed at the beginning of testing. While this reduces latency, it comes at the expense of (potentially) more used tests. Both variants have applications depending on the more expensive resource (testing time vs. the amount of tests), whereas the former is the more intuitive variant of group testing and was also used for example in pooled PCR tests for SARS-CoV-2 Mutesa et al. (2021); Singh et al. (2020).

Our work provides a rigorous and complete understanding of both adaptive and non-adaptive group testing in the aforementioned noisy setting for the linear regime. To be precise, we provide exact thresholds on the number of tests such that exact recovery is impossible below the threshold, and possible above the threshold. Moreover, we provide efficient algorithms for the latter case.

1.1. Contribution

We consider group testing under the *i.i.d. prior*, where the n individuals are infected independently with constant probability α,[1] with test results being observed through a noisy

1. Note that this can be transformed to the other commonly studied studied prior, the *combinatorial prior*, where the number of infected individuals is fixed; cf. Appendix to Chapter 1 in Aldridge et al. (2019).



(a) Symmetric noise (b) Asymmetric noise

Figure 1: Adaptive and non-adaptive thresholds $c_{\mathrm{ad}} = m_{\mathrm{ad}}/(n \ln n)$ and $c_{\mathrm{na}} = m_{\mathrm{na}}/(n \ln n)$.

binary channel. For both adaptive and non-adaptive schemes, we pin down the number $m = cn\ln(n)$ of necessary tests down to the constant c. We call this threshold m_{na} for the non-adaptive and m_{ad} for the adaptive case. To prove these thresholds, we pinpoint the precise combinatorial conditions necessary for the exact identification of infected individuals. Furthermore, we harness this knowledge to provide *efficient* algorithms for both settings using $(1+\varepsilon)m_{\mathrm{na}}$ (resp., $(1+\varepsilon)m_{\mathrm{ad}}$) tests.

We now give a brief overview of the main obstacles for estimating an individual's infection status. In the non-adaptive case, each individual requires tests whose results change when the individual's status changes; call such tests *good* (for said individual). With a fixed test design, the number of good tests fluctuates depending on the random infection statuses. While there is no fixed lower bound on the necessary number of good tests required for *every single* individual, there need to be sufficiently many tests so that there are not too many individuals that don't have enough good tests. If not, even estimators having oracle access to the ground truth except for the individual in question (*genie-based estimators*) fail.

In the adaptive case, on a conceptual level, there are two important differences from the non-adaptive case: First, an adaptive test design can ensure that most tests are used on infected individuals. Second, an adaptive test design can ensure that almost all of these tests are good for some (infected) individuals, removing the fluctuation of the non-adaptive case. In this sense, the main barrier is that all infected individuals need to participate in enough tests such that there are enough good tests to exclude the possibility of the individual *not* being infected after all.

Our efficient algorithms, the non-adaptive `SPOG` (***s**ynthetic **p**seud**o**-**g**enie*) and the three-stage adaptive `PRESTO` (***pre**-**s**orting **t**hresh**o**lder*), overcome these obstacles with a test count just above the thresholds. `SPOG` uses almost all its tests on random groups of an optimal size Γ depending on both α and a measure of noisiness. Few individual tests are used to construct a "synthetic pseudo-genie"; `SPOG`'s output is based on the group tests which are good by the pseudo-genie. Similarly, `PRESTO`'s first stage uses a few individual tests to pre-sort individuals into a mostly-infected and a mostly-uninfected set. Its second stage weeds out all uninfected individuals among the former by thresholding on many individual tests. The final stage uses `SPOG` twice to identify the few infected individuals among the first stage's mostly uninfected set, as well as the second stage's rejects. We note that both algorithms classify almost all individuals correctly as pre-processing based on only a negligible fraction of all tests. Hence, approximate recovery requires significantly fewer tests in both settings; determining the exact number of tests required for this remains open.

Figure 1 depicts the thresholds' constants c for different values of α and different channels. It is blatant how adaptivity saves a significant amount of tests in all depicted scenarios: in Figure 1(a), we see that for a symmetric channel, the adaptive algorithm uses fewer tests for even tenfold higher noise. Additionally, viewing the graphs of the non-adaptive threshold, we see that they do not seem to be differentiable for all α. Indeed, these "kinks" appear due to the optimization over the (discrete) test degrees. On another note, Figure 1(b) shows two additional interesting properties: First, the non-adaptive threshold is not sensitive to switching p_{01} and p_{10}, as the blue and cyan lines overlap. On the other hand, in the adaptive case, false positive results require more tests than false negatives.

1.2. Related work

In recent years, there has been a variety of work on group testing and its variants. An excellent overview of results prior to 2019 can be found in Aldridge et al. (2019). As mentioned above, the overwhelming majority of recent rigorous results focus on the case of a sublinear number of infected individuals. Following a considerable amount of prior work Aldridge et al. (2014); Scarlett and Cevher (2016); Johnson et al. (2018); Coja-Oghlan et al. (2020), in the sublinear regime, noiseless group testing is now understood completely Coja-Oghlan et al. (2021), while noisy non-adaptive group testing is understood almost completely Coja-Oghlan et al. (2024); Chen and Scarlett (2024). However, results for the linear regime are much more limited. In the noiseless case, one cannot improve on individual testing once the fraction of infected individual exceeds $\alpha = \frac{1}{2}(3-\sqrt{5})$ Ungar (1960). Below this threshold, one has to differentiate between non-adaptive and adaptive group testing. For non-adaptive group testing, one cannot do better than individual testing either Aldridge (2019a). In the adaptive case important contributions are an algorithm Aldridge (2019b) based on binary splitting Hwang (1972) as well as Aldridge (2020) that introduces a conservative two stage adaptive test scheme and lower bounds for two stage testing. However, determining a tight threshold for noiseless linear adaptive group testing remains open.

If we turn from noiseless to noisy group testing rigorous results are even more sparse. To the best of our knowledge, the only result in this setting is a lower bound on the necessary tests for the special case of symmetric noise Scarlett (2019), which is a factor $(1-2p_{01})^{-1}$ smaller than the actual bound. Indeed, our proposed adaptive test scheme is conceptually similar to the three-stage test scheme for symmetric noise of Scarlett (2019), which is analyzed only for the sublinear regime. However, we are able to generalize this approach to general noise channels and prove that our test scheme uses the optimal number of tests in the linear regime.

The present work's proofs are partially inspired by techniques from community detection such as focusing on a small portion of individuals that are already hard to estimate Abbe and Sandon (2015), as well as the use of a genie-based estimator for impossibility proofs Abbe (2017). The latter is, e.g., also used in work on computation with noisy queries (cf. Feige et al. (1994)).

2. Model and Results

Model and commonly used notation We consider group testing under the *i.i.d. prior*, where the infection statuses of all of the n individuals are independent, with each individual's

infection probability being $\alpha \in (0,1)$. We call a $\sigma \in \{0,1\}^n$ an *infection vector*, where $\sigma(i) = 1$ if i is infected under σ, and $\sigma(i) = 0$ otherwise. We then write $\boldsymbol{\sigma}$ for the random infection vector (the *ground truth*) as described above. Often, we consider a test design as a bipartite graph $G = (V, F, E)$, with $V = V(G) = [n]$ being the set of individuals, $F = F(G)$ being the set of tests, and the set E consisting of the edges between individuals and tests, with an edge being present if and only if a given test contains a given individual. We write $\mathfrak{G}_m$ for the set of all such graphs with m tests.

For a test $a \in F$, we write $\partial_G a$ for the set of individuals contained in a, for an individual $i \in [n]$, we write $\partial_G i$ for the set of tests containing i. When G is clear from context (as is the case most of the time), we omit the subscript. If the tests were perfectly accurate, then a test would appear positively if and only if it contains an infected individual. So writing $\boldsymbol{\tau} = (\boldsymbol{\tau}(a))_{a\in F} \in \{0,1\}^F$ for the vector of these hypothetical test results, we have $\boldsymbol{\tau}(a) = \max\{\boldsymbol{\sigma}(i) \mid i \in \partial a\}$. However, we consider the test results as being transmitted through a noisy channel: We write $\tilde{\boldsymbol{\tau}}$ for the vector of observed test results. $\tilde{\boldsymbol{\tau}}(a)$ depends only on $\boldsymbol{\tau}(a)$, with the channel noise being independent for all tests. Often, for a set of tests $S \subseteq F$, we write $S^+ = \{a \in S \mid \tilde{\boldsymbol{\tau}}(a) = 1\}$ for the subset of S with positive observed test result, and $S^- = S \setminus S^+$ for the subset of S with negative observed test result.

For $k, l \in \{0,1\}$, we let $p_{kl} \in (0,1)$ be the probability that a test with hypothetical test result $\boldsymbol{\tau}(a) = k$ displays as l. So formally, for any $\tau, \tilde{\tau} \in \{0,1\}^m$, $\mathbb{P}[\tilde{\boldsymbol{\tau}} = \tilde{\tau} \,|\, \boldsymbol{\tau} = \tau] = \prod_{j\in[m]} p_{\tau_j \tilde{\tau}_j}$. We write $\boldsymbol{p} = (p_{00}, p_{01}, p_{10}, p_{11})$ for the vector of these probabilities, and call $\boldsymbol{p}$ a *noisy channel*. Note that we must have $p_{00} + p_{01} = p_{10} + p_{11} = 1$, that $p_{11} = 1 - p_{10}$ is the *sensitivity* (true positive rate) of the test, and that $p_{00} = 1 - p_{01}$ is the *specificity* (true negative rate) of the test. Additionally, we require that $p_{01} + p_{10} < 1$: if $p_{01} + p_{10} = 1$, then tests containing, or not containing, an infected individual would be indistinguishable (the channel has capacity 0); and if $p_{01} + p_{10} > 1$, one can flip all observed test results to obtain a channel with $p_{01} + p_{10} < 1$. We call a noisy channel satisfying these conditions *valid*. For $p, q \in [0,1]$, we write $\mathrm{D}_{\mathrm{KL}}(p\|q)$ for the Kullback–Leibler divergence (or relative entropy) of $\mathrm{Ber}(p)$ from $\mathrm{Ber}(q)$, i.e., $\mathrm{D}_{\mathrm{KL}}(p\|q) = p \ln \frac{p}{q} + (1-p) \ln \frac{1-p}{1-q}$. Furthermore, for any valid noisy channel $\boldsymbol{p}$, define

$$\beta = \beta(\boldsymbol{p}) = \max_{c \in [p_{01}, p_{11}]} \min\left\{\mathrm{D}_{\mathrm{KL}}(c\|p_{01}), \mathrm{D}_{\mathrm{KL}}(c\|p_{11})\right\}, \tag{1}$$

and let $C = C(\boldsymbol{p})$ be that c where the maximum is obtained. Note that as $p_{01}, p_{11} \in (0,1)$, we have $\beta(\boldsymbol{p}) = \mathrm{D}_{\mathrm{KL}}(C(\boldsymbol{p})\|p_{01}) = \mathrm{D}_{\mathrm{KL}}(C(\boldsymbol{p})\|p_{11})$, and $0 < \beta(\boldsymbol{p}) < \infty$ for any valid noisy channel $\boldsymbol{p}$ (see Lemma C.1 of the full version of this paper Hintze et al. (2024) for a proof). Finally, we say that an event occurs *with high probability* (w.h.p.) if it occurs with probability $1 - o_n(1)$.

Results For the non-adaptive case, we show in Section 4.1 that for

$$m_{\mathrm{na}} = m_{\mathrm{na}}(\alpha, \boldsymbol{p}) = \min_{\Gamma \in \mathbb{N}^+} \frac{n \ln n}{-\Gamma \cdot \ln\left(1 - (1-\alpha)^{\Gamma-1} \cdot \left(1 - \mathrm{e}^{-\beta(\boldsymbol{p})}\right)\right)} \tag{2}$$

when $(1-\varepsilon)m_{\mathrm{na}}$ tests are used, any non-adaptive algorithm will fail with high probability:

Theorem 1 *For any valid noisy channel $\boldsymbol{p}$, $\alpha > 0$ and $\varepsilon > 0$, there exists some $n_0 = n_0(\boldsymbol{p}, \alpha, \varepsilon)$ such that for every $n > n_0$, all test designs G with $m < (1-\varepsilon)m_{\mathrm{na}}(\alpha, \boldsymbol{p})$ tests*

and every estimation function $f_G : \{0,1\}^m \to \{0,1\}^n$*:*

$$\mathbb{P}[f_G(\tilde{\boldsymbol{\tau}}_G) = \boldsymbol{\sigma}] \leq \varepsilon.$$

Furthermore, we show that the non-adaptive algorithm `SPOG` given in Section 4.2 recovers $\boldsymbol{\sigma}$ with high probability using $(1+\varepsilon)m_{\mathrm{na}}$ tests.

Theorem 2 *For any valid noisy channel* $\boldsymbol{p}$*,* $\alpha > 0$ *and* $\varepsilon > 0$*, there exists some* $n_0 = n_0(\boldsymbol{p}, \alpha, \varepsilon)$ *such that for every* $n > n_0$*, there is a randomized test design* $\boldsymbol{G}$ *using* $m \leq (1+\varepsilon)m_{\mathrm{na}}(\alpha, \boldsymbol{p})$ *tests w.h.p. and a deterministic polynomial time algorithm* `SPOG` *such that*

$$\mathbb{P}[\texttt{SPOG}(\boldsymbol{G}, \tilde{\boldsymbol{\tau}}_{\boldsymbol{G}}) = \boldsymbol{\sigma}] \geq 1 - \varepsilon.$$

Our second pair of results states that

$$m_{\mathrm{ad}} = m_{\mathrm{ad}}(\alpha, \boldsymbol{p}) = \frac{\alpha}{\mathrm{D}_{\mathrm{KL}}\left(p_{11} \| p_{01}\right)} \cdot n \ln n$$

is the same kind of threshold for adaptive schemes. In Section 5, we see that when only $(1-\varepsilon)m_{\mathrm{ad}}$ tests are used, any adaptive algorithm fails with high probability:

Theorem 3 *For any valid noisy channel* $\boldsymbol{p}$*,* $\alpha > 0$ *and* $\varepsilon > 0$*, there exists some* $n_0 = n_0(\boldsymbol{p}, \alpha, \varepsilon)$ *such that for every* $n > n_0$*, any adaptive test scheme* $\boldsymbol{\mathcal{A}}$ *using* $m_{\boldsymbol{\mathcal{A}}} < (1-\varepsilon)m_{\mathrm{ad}}(\alpha, \boldsymbol{p})$*, and any estimation algorithm* $f_{G_{\boldsymbol{\mathcal{A}}}} : \{0,1\}^{m_{\boldsymbol{\mathcal{A}}}} \to \{0,1\}^n$*:*

$$\mathbb{P}[f_{G_{\boldsymbol{\mathcal{A}}}}(\tilde{\boldsymbol{\tau}}_{\boldsymbol{\mathcal{A}}}) = \boldsymbol{\sigma}] \leq \varepsilon. \tag{3}$$

Furthermore we show that the adaptive algorithm `PRESTO` given in Section 4.2 recovers $\boldsymbol{\sigma}$ with high probability using $(1+\varepsilon)m_{\mathrm{ad}}$ tests.

Theorem 4 *For any valid noisy channel* $\boldsymbol{p}$*,* $\alpha > 0$ *and* $\varepsilon > 0$*, there exists some* $n_0 = n_0(\boldsymbol{p}, \alpha, \varepsilon)$ *such that for every* $n > n_0$*, the three-stage adaptive test scheme* `PRESTO` *uses at most* $m \leq (1+\varepsilon)m_{\mathrm{ad}}(\alpha, \boldsymbol{p})$ *w.h.p. such that*

$$\mathbb{P}[\texttt{PRESTO}(\tilde{\boldsymbol{\tau}}_{\texttt{PRESTO}}) = \boldsymbol{\sigma}] \geq 1 - \varepsilon. \tag{4}$$

3. Preliminaries

MAP estimation For the impossibility results we have to prove that *any* estimator fails w.h.p. if not supplied with enough tests. Since dealing with an arbitrary estimator is tedious, instead we first consider the estimator maximizing the probability of exactly recovering $\boldsymbol{\sigma}$, namely the maximum a posteriori (MAP) estimator, which chooses the infection vector that maximizes the a posteriori probability given observed test results.

Definition 5 (MAP estimate) *For any test design* $G \in \mathfrak{G}_m$ *and any observed results* $\tilde{\tau} \in \{0,1\}^m$*, the* MAP estimate $\hat{\sigma}_{\mathrm{MAP}}(G, \tilde{\tau}) = \hat{\sigma}^{g,\tilde{\tau}}_{\mathrm{MAP}}$ *of* $\boldsymbol{\sigma}$ *is given by*

$$\hat{\sigma}^{G,\tilde{\tau}}_{\mathrm{MAP}} = \operatorname*{arg\,max}_{\hat{\sigma} \in \{0,1\}^n} \mathbb{P}[\hat{\sigma} = \boldsymbol{\sigma} \mid G, \tilde{\tau}], \tag{5}$$

where, if ambiguous, we choose an arbitrary $\hat{\sigma}$ *that maximizes also the number of zeroes contained.*

As is commonly known, no estimator is better than the MAP estimator for achieving exact recovery (see, e.g., (Abbe, 2017, Section 3.2)).

Notation used throughout For brevity, for any infection vector $\sigma \in \{0,1\}^n$, we write $\mathcal{I}_\sigma$ for the set of infected individuals under σ, i.e., $\mathcal{I}_\sigma = \{i \in [n] \mid \sigma(i) = 1\}$, and hence $\overline{\mathcal{I}_\sigma} = [n] \setminus \mathcal{I}_\sigma$ is the set of uninfected individuals. Consequently, $\mathcal{I}_{\boldsymbol{\sigma}}$ is the (random) set of infected individuals.

Furthermore, throughout our paper, we distinguish between tests that are *good* for estimating an individual and those that are not. This differentiation of tests has been commonly used in group testing (see, e.g., Coja-Oghlan et al. (2024); Aldridge (2019a)). Intuitively, a good test for an individual actually carries information about the individual's infection status, which is only the case if no other infected individual is included.

Definition 6 (good tests) *For a given test design G, a test a is* good *for an individual i under the infection vector $\sigma \in \{0,1\}^i$ if a contains i, and no individual in a, except for (perhaps) i, is infected. We write $g_i(\sigma, G)$ for the number of such tests for i under σ. For a given vector $\tilde{\tau}$ of observed test results, we also write $g_i^-(\sigma, G, \tilde{\tau})$ (resp., $g_i^+(\sigma, G, \tilde{\tau})$) for the number of such tests displaying negatively (resp., positively). When σ is clear from context, we may simply write g_i (etc.). Formally,*

$$\begin{aligned} g_i(\sigma, G) &= |\{a \in \partial i \mid (\partial a \setminus \{i\}) \cap \mathcal{I}_\sigma = \emptyset\}|\,, \\ g_i^-(\sigma, G, \tilde{\boldsymbol{\tau}}) &= |\{a \in \partial i \mid \tilde{\boldsymbol{\tau}}(a) = 0 \wedge (\partial a \setminus \{i\}) \cap \mathcal{I}_\sigma = \emptyset\}| \text{ and} \\ g_i^+(\sigma, G, \tilde{\boldsymbol{\tau}}) &= g_i(\sigma, G) - g_i^-(\sigma, G, \tilde{\boldsymbol{\tau}})\,. \end{aligned}$$

As in other places, we omit the parameters when the values are clear from context.

4. Non-adaptive Group Testing

In this section, we outline the proof strategy for both impossibility as well as achievability in the non-adaptive setting. In both, we (conceptually) use an estimator which is even better than the MAP estimator, the so-called *genie-based estimator*. This technique is inherited from community detection Abbe (2017); Abbe and Sandon (2015) but so far not used in the context of group testing to the best of our knowledge.

Definition 7 (Genie estimator) *Given a test design G and the observed results $\tilde{\tau}$, with $\boldsymbol{\sigma}_{-i}$ the ground truth for every other individual except i, the* genie-based estimator *(or just* genie estimator*) is given by*

$$\hat{\sigma}_{\text{gen}}^{G,\tilde{\tau}}(i) = \hat{\sigma}_{\text{gen}}^{G,\tilde{\tau},\boldsymbol{\sigma}_{-i}}(i) = \operatorname*{arg\,max}_{s \in \{0,1\}} \mathbb{P}[\boldsymbol{\sigma}(i) = s \,|\, G, \tilde{\boldsymbol{\tau}} = \tilde{\tau}, \boldsymbol{\sigma}_{-i}]. \tag{6}$$

If both 0 and 1 maximize $\mathbb{P}[\boldsymbol{\sigma}(i) = s \,|\, G, \tilde{\boldsymbol{\tau}} = \tilde{\tau}, \boldsymbol{\sigma}_{-i}]$, let $\hat{\sigma}_{\text{gen}}^{G,\tilde{\tau}}(i) = 0$.

Intuitively, it is clear that the genie estimator outperforms the MAP estimator: it has access not only to the test design and the observed results but also to the whole ground truth $\boldsymbol{\sigma}_{-i}$ to determine $\boldsymbol{\sigma}(i)$. So proving that even the genie estimator fails when the test design contains too few tests—as done in Section 4.1—implies that all estimators fail. The following lemma, proven in Appendix D.0 of the full version Hintze et al. (2024), formalizes this intuition.

Lemma 8 (Genie estimator better than MAP estimator) *For every test design G,*

$$\mathbb{P}[\boldsymbol{\sigma} = \hat{\sigma}^{G,\tilde{\boldsymbol{\tau}}}_{\text{gen}}] \geq \mathbb{P}[\boldsymbol{\sigma} = \hat{\sigma}^{G,\tilde{\boldsymbol{\tau}}}_{\text{MAP}}].$$

Of course, the genie estimator is not realizable: when classifying a given individual, it requires oracle access to the true infection status of all other individuals. However, our algorithm `SPOG` presented in Section 4.2 uses the concept by emulating the oracle access well enough for the classification to still be correct w.h.p.

4.1. Impossibility: proof strategy for Theorem 1

For the lower bound in the non-adaptive case, we consider an arbitrary test design G with fewer than $(1-\varepsilon)m_{\text{na}}$ tests. First, we modify the test design to achieve some properties crucial to our analysis without decreasing the success probability of the genie estimator significantly, if at all. We identify individuals that are especially hard to estimate for the genie estimator when using G. Then, we show that there must be a large set of such individuals with independent tests, such that the genie estimator fails w.h.p.

To execute this plan, we first need to characterize the genie estimator. Observe that its estimate for an individual i depends only on the observed results of good tests for i, i.e., tests that do not contain any other infected individual. The following lemma shows that the genie estimator is equivalent to a thresholding function on the fraction of positively displayed good tests. A proof can be found in Appendix D.1.1 of the full version Hintze et al. (2024). Note that the C defined in the lemma statement is equal to the maximizer of (1) as provided by Lemma C.1 of the full version Hintze et al. (2024).

Lemma 9 *Let $C = \ln(\frac{p_{00}}{p_{10}})/\ln(\frac{p_{11}p_{00}}{p_{01}p_{10}})$, and $\kappa = \kappa(\alpha, \boldsymbol{p}) = \ln(\frac{\alpha}{1-\alpha})/\ln(\frac{p_{01}p_{10}}{p_{11}p_{00}})$. Then*

$$\hat{\sigma}^{G}_{\text{gen}}(i) = \begin{cases} 0 & \text{if } g_i(\boldsymbol{\sigma}) = 0 \text{ and } \alpha \leq \frac{1}{2}, \text{ or } g_i^+(\boldsymbol{\sigma}) \leq C g_i(\boldsymbol{\sigma}) + \kappa, \\ 1 & \text{otherwise.} \end{cases}$$

We now describe our modifications to the test design. First, we limit the degrees of both individuals and tests in the modified design. This enables us to find a sufficiently large set of individuals such that their tests, and hence their genie estimates, are independent. As for tests, we show the following in Appendix D.1.2 of the full version Hintze et al. (2024):

Lemma 10 *If $m = \mathcal{O}(n\ln(n))$, then w.h.p., all tests a with $|\partial a| \geq \ln^2(n)$ contain at least two infected individuals.*

Consequently, all such tests are not good for *any* individual w.h.p., so they are not considered by the genie estimator, Hence, they are safe to remove without affecting the genie estimator's success probability too much. As for the individuals, we remove all which participate in too many tests, and assume that they are correctly estimated. This cannot decrease the genie estimator's success probability. Additionally, for each individual, we add $\eta\ln(n)$ individual tests for some small $\eta > 0$ to be determined later. This enforces a lower bound on the number of good tests for each individual, which we use in a few calculations. To summarize:

Definition 11 *Given a test design G, the* modified test design G_η *is constructed as follows:*

- *remove tests in* $L = \left\{a \in F \mid |\partial_G a| \geq \ln^2(n)\right\}$ *to obtain design* G',
- *remove individuals in* $J = \{i \in [n] \mid |\partial_{G'} i| > \ln^4(n) - \lceil \eta \ln(n) \rceil\}$ *to obtain design* G'', *and*
- *add* $\lfloor \eta \ln(n) \rfloor$ *individual tests for each individual* i *with* $\eta > 0$ *to obtain design* G_η.

Note that once we remove individuals with a large degree we still have

$$\check{n} = n - cn \ln(n) \ln^2(n) / \ln^4(n) = n - cn / \ln(n) = n(1 - o(1))$$

individuals left to be estimated. W.l.o.g., we assume that these are the first $\check{n}$ individuals $[\check{n}]$. The following lemma confirms the intuition that the genie estimator can only be improved by the adjustments in the modified test design; we prove this in Appendix D.1.2 of the full version Hintze et al. (2024).

Lemma 12 (modified test setup is easier) *Let* $\boldsymbol{\sigma}[\check{n}]$ *be the infection statuses for individuals in* $[\check{n}]$ *and* $\hat{\sigma}[\check{n}]$ *be the prediction of the infection statuses for individuals in* $[\check{n}]$ *based on estimator* $\hat{\sigma}$. *Then*

$$\mathbb{P}\left[\boldsymbol{\sigma}[\check{n}] = \hat{\sigma}_{\text{gen}}^{G_\eta}[\check{n}]\right] \geq \mathbb{P}\left[\boldsymbol{\sigma}[\check{n}] = \hat{\sigma}_{\text{MAP}}^{G_\eta}[\check{n}]\right] \geq \mathbb{P}\left[\boldsymbol{\sigma} = \hat{\sigma}_{\text{MAP}}^{G'}\right] \geq \mathbb{P}\left[\boldsymbol{\sigma} = \hat{\sigma}_{\text{MAP}}^{G}\right] - o(1)\,. \quad (7)$$

Now in this modified test design, we find a relatively large set of individuals whose tests are independent. This is ensured by not including individuals whose tests intersect, i.e., by making sure that their second neighborhoods are disjoint. We call such sets of individuals *distant*. We choose such a set greedily by repeatedly selecting an individual i (by some arbitrary criterion) and removing every individual of distance at most four from i in G_η. Since there are at most $(\ln^4(n) \cdot \ln^2(n))^2 = \ln^{12}(n)$ removed in every iteration (by our limit on the respective degrees), we can repeat this at least $\check{n} / \ln^{13}(n)$ times without running out of individuals.

Our goal is to show that in any modified test design we can choose a distant set such that w.h.p. at least one individual in this set is misclassified. To that end, for any (distant) set of individuals D, let $\boldsymbol{D}_{\text{err}}$ be the set of those individuals in D that are misclassified by the genie estimator on G_η, i.e., $\boldsymbol{D}_{\text{err}} = \{i \in D \mid \hat{\sigma}_{\text{gen}}^{G_\eta}[\check{n}](i) \neq \boldsymbol{\sigma}[\check{n}](i)\}$. We first bound the probability that the genie estimator is correct conditioned on $\boldsymbol{\sigma}$. The proof in Appendix D.1.3 of the full version Hintze et al. (2024) exploits the fact that correctness for all individuals requires correctness for any given distant set of individuals. Then, we apply Chebyshev's inequality to $|\boldsymbol{D}_{\text{err}}|$ using the independence of distant individuals' tests.

Lemma 13 *For any set of distant individuals* $D \subseteq [\check{n}]$, $\mathbb{P}[\boldsymbol{\sigma}[\check{n}] = \hat{\sigma}_{\text{gen}}^{G_\eta}[\check{n}] \mid \boldsymbol{\sigma}] \leq (\mathbb{E}[|\boldsymbol{D}_{\text{err}}| \mid \boldsymbol{\sigma}])^{-1}$.

To bound $\mathbb{E}[|\boldsymbol{D}_{\text{err}}| \mid \boldsymbol{\sigma}]$ from below, we obtain the following bound on the probability that any specific individual in D is misclassified by the genie estimator (conditioned on $\boldsymbol{\sigma}$) in Appendix D.1.4 of the full version Hintze et al. (2024).

Lemma 14 (probability of misclassification) *Let* $i \in [\check{n}]$, *let* $C = \ln(\frac{p_{00}}{p_{10}}) / \ln(\frac{p_{11}p_{00}}{p_{01}p_{10}})$ *and let* $0 < \delta < 1 - p_{10} - C$. *For sufficiently large* n *the probability that the genie estimator misclassifies* i *is bounded from below as*

$$\mathbb{P}\left[\hat{\sigma}_{\text{gen}}^{G_\eta}(i) \neq \boldsymbol{\sigma}(i) \,\middle|\, \boldsymbol{\sigma}\right] \geq \exp\left(-g_i(\boldsymbol{\sigma}, G_\eta)\, \mathrm{D}_{\mathrm{KL}}\left(C - \delta \| p_{11}\right)\right). \quad (8)$$

All that remains is to find a distant set D such that the sum of the lower bounds given by Lemma 14 over all $i \in D$ is large for most $\boldsymbol{\sigma}$. The following lemma achieves this.

Lemma 15 *$\forall \varepsilon, \delta'' > 0$ $\exists \delta, \eta > 0$ $\exists n_0(\delta'')$ $\forall n \geq n_0(\delta'')$, any test design G using at most $(1-\varepsilon)m_{\mathrm{na}}$ tests, there is a distant set $D \subseteq [\check{n}]$ so that*

$$\mathbb{P}\left[\sum_{i \in D} \exp\left(-\operatorname{D_{KL}}\left(C - \delta \| p_{11}\right) g_i(\boldsymbol{\sigma}, G_\eta)\right) > \frac{1}{2\delta''}\right] \geq 1 - 4\delta''. \tag{9}$$

The proof in Appendix D.1.5 of the full version Hintze et al. (2024) exploits the fact that for any distant set D, the sum in question is a sum of $|D|$ independent and bounded random variables. So it is well-concentrated, and it suffices to bound it from below in expectation. That bound in turn hinges on the fact that the events of tests being good for individuals are monotonic in $\boldsymbol{\sigma}$ which enables the application of the FKG inequality. Using the fact that a test a is good for an individual with probability $(1-\alpha)^{|\partial a|-1}$, some further calculation shows that, writing $b = \operatorname{D_{KL}}\left(C - \delta \| p_{11}\right)$,

$$\ln\left(\mathbb{E}[\exp(-b g_i(\boldsymbol{\sigma}, G_\eta))]\right) \geq \sum_{a \in \partial i} \ln(1 - (1-\alpha)^{|\partial a|-1}(1 - e^{-b})).$$

When choosing the distant set D greedily to maximize the right-hand side, the average of this over D is at worst just slightly below the average over *all* individuals. This average comes down to

$$\frac{1}{n} \sum_{i \in [n]} \sum_{a \in \partial i} \ln(1 - (1-\alpha)^{|\partial a|-1}(1 - e^{-b})) = \frac{1}{n} \sum_{a \in F(G_\eta)} |\partial a| \ln(1 - (1-\alpha)^{|\partial a|-1}(1 - e^{-b})).$$

For $|\partial a| = \Gamma$, and $b \overset{\delta \to 0}{\to} \beta$, this is the value optimized over in the definition of the threshold m_{na}. Now we plug everything together to conclude the proof of Theorem 1.

Proof [Proof of Theorem 1] For sufficient large n, the $o(1)$ term in Lemma 12 is at most $\varepsilon/2$. It now suffices to show that the probability of the genie estimator being correct on G_η is at most $\varepsilon/2$ for sufficiently large n. To bound the probability of the genie estimator being correct on $[\check{n}]$, combining Lemmas 13 and 14, for any set of distant individuals $D \subseteq [\check{n}]$,

$$\mathbb{P}\left[\boldsymbol{\sigma}[\check{n}] = \hat{\sigma}_{\mathrm{gen}}^{G_\eta}[\check{n}] \,\middle|\, \boldsymbol{\sigma}\right] \leq \left(\sum_{i \in D} \exp\left(-g_i(\boldsymbol{\sigma}, G_\eta) \operatorname{D_{KL}}\left(C - \delta \| p_{11}\right)\right)\right)^{-1}. \tag{10}$$

Now write $\boldsymbol{U}_D$ for this upper bound. By Lemma 15, for any δ'' and sufficiently small δ'', we can choose D so that $\mathbb{P}[\boldsymbol{U}_D \leq 2\delta''] \geq 1 - 4\delta''$. Since $\mathbb{P}[\boldsymbol{\sigma}[\check{n}] = \hat{\sigma}_{\mathrm{gen}}^{G_\eta}[\check{n}] \mid \boldsymbol{\sigma}] \leq 1$, we have

$$\begin{aligned}
\mathbb{P}\left[\boldsymbol{\sigma}[\check{n}] = \hat{\sigma}_{\mathrm{gen}}^{G_\eta}[\check{n}]\right] &= \mathbb{E}\left[\mathbb{P}\left[\boldsymbol{\sigma}[\check{n}] = \hat{\sigma}_{\mathrm{gen}}^{G_\eta}[\check{n}] \,\middle|\, \boldsymbol{\sigma}\right]\right] \leq \mathbb{E}[\min\{1, \boldsymbol{U}_D\}] && (11)\\
&\leq \mathbb{E}\left[\min\{1, \boldsymbol{U}_D\} \,\middle|\, \boldsymbol{U}_D \leq 2\delta''\right] + \mathbb{P}[\boldsymbol{U}_D > 2\delta''] = 2\delta'' + 4\delta'', && (12)
\end{aligned}$$

which yields the theorem with $\delta'' = \varepsilon/12$. ∎

4.2. The algorithm SPOG: proof strategy for Theorem 2

We now introduce our non-adaptive test design as well as the corresponding estimation algorithm and prove that it satisfies Theorem 2. Conceptually, the algorithm performs classification in two stages—however, all the tests are determined in advance, so it is indeed a non-adaptive algorithm.

In the first stage, $\lceil \eta \ln(n) \rceil$ individual tests per individual (for some small $\eta = o(1) > 0$, precisely determined later) are used for pre-classification: $\hat{\boldsymbol{\sigma}}^{(1)}$ classifies an individual as infected if at least a C proportion of these individual tests display positively, where C is the threshold defined below (1). This then acts as a synthetic "pseudo-genie" in the second stage: it essentially behaves similarly to a genie estimator, with the always-correct genie being replaced by the mostly-correct "pseudo-genie" created in the first stage.

The second stage's tests are almost entirely group tests all having the same size Γ, with some extra individual tests to guarantee that each individual participates in sufficiently many tests. The classification is not based on *all* second-stage tests containing a given individual, but rather just a subset: First, the algorithm constructs a set of tests D_i such that any two tests in D_i only intersect in i. We call such sets *distinctive*. For each variable i, SPOG finds such a set D_i by keeping track of the "already used" second neighborhood S. These sets have the useful property that their observed results are independent when conditioning on $\boldsymbol{\sigma}(i)$. Then, SPOG only considers tests in D_i which are good under the pseudo-genie $\hat{\boldsymbol{\sigma}}^{(1)}$; this is the set of *pseudo-good* tests P_i.

To describe the design $\boldsymbol{G}_{\texttt{SPOG}}$ let $\xi = \xi(\alpha, \boldsymbol{p}, \Gamma) = \left(-\ln\left(1 - (1-\alpha)^{\Gamma-1} \cdot (1 - \mathrm{e}^{-\beta(\boldsymbol{p})})\right)\right)^{-1}$. This ensures that an individual participating in $\xi \ln(n)$ tests of size Γ is misclassified by the genie estimator with probability $\mathcal{O}(n^{-1})$.

Definition 16 (Test design for SPOG) *For $\hat{\alpha} \in (0,1)$, $\Gamma \in \mathbb{N}^+$ with $\Gamma \leq \lceil \hat{\alpha}^{-1} \rceil$, a valid noisy channel $\boldsymbol{p}$, and $\eta, \varepsilon > 0$, we let $\boldsymbol{G}_{\mathit{SPOG}}(n, \hat{\alpha}, \boldsymbol{p}, \Gamma, \eta, \varepsilon)$ be the test design having n individuals and with its test set F being the union of*

- *the set F_1 of $\lceil \eta \ln n \rceil$ individual tests $F_1(i)$ per individual $i \in [n]$,*
- *the set $F_{2,grp}$ of $\lceil (1+\varepsilon/3)\xi \ln n/\Gamma \rceil$ group tests chosen i.u.r. among tests of size Γ, and*
- *the set $F_{2,idv}$ of $\max\left\{0, \left(1+\frac{\varepsilon}{6}\right)\xi \ln n + 1 - |F_{2,grp} \cap \partial i|\right\}$ individual tests for each $i \in [n]$.*

Furthermore, we write $F_2 = F_{2,grp} \cup F_{2,idv}$.

Our analysis makes use of the following bounds on ξ proven in Appendix D.2.1 of the full version Hintze et al. (2024). By part *(b)*, the Γ minimizing the number of tests satisfies the restriction $\Gamma \leq \lceil \alpha^{-1} \rceil$, so that this was safe after all.

Lemma 17 *Let $\alpha \in (0,1)$, and $\boldsymbol{p}$ a valid noisy channel.*

(a) For all $\Gamma \leq \lceil \alpha^{-1} \rceil$, $\xi(\alpha, \boldsymbol{p}, \Gamma) \in [\beta^{-1}, \frac{\mathrm{e}}{(1-\alpha)(1-\mathrm{e}^{-\beta})}]$, and $\xi(\alpha, \boldsymbol{p}, \Gamma)$ is increasing in α.

(b) Let $\hat{\Gamma} \in \arg\max_{\Gamma \in \mathbb{N}^+} -\Gamma \ln\left(1 - (1-\alpha)^{\Gamma-1}\left(1 - \mathrm{e}^{-\beta(\boldsymbol{p})}\right)\right)$. Then $\hat{\Gamma} \leq \lceil \alpha^{-1} \rceil$.

The following lemma bounds the number of tests in $G_{\texttt{SPOG}}$. The crux of the proof in Appendix D.2.2 of the full version Hintze et al. (2024) is that the probability of adding individual tests via $F_{2,idv}$ for a given i is polynomially small. Hence, even if $\Theta(\ln(n))$ tests were added for any such individual, $|F_{2,idv}|$ is negligible.

Data: Instance G of $\boldsymbol{G}_{\texttt{SPOG}}$, $\tilde{\boldsymbol{\tau}}_G \in \{0,1\}^n$
for $i \in [n]$ **do**
  $\hat{\boldsymbol{\sigma}}^{(1)}(i) \leftarrow \mathbb{1}\left[\left|F_1^+(i)\right| \geq C \cdot |F_1(i)|\right]$
for $i \in [n]$ **do**
  $S \leftarrow \emptyset$, $D_i \leftarrow \emptyset$
  for $a \in F_2 \cap \partial i$ **do**
    if $\partial a \setminus \{i\} \cap S = \emptyset$ **then**
      $D_i \leftarrow D_i \cup \{a\}$, $S \leftarrow S \cup (\partial a \setminus \{i\})$
  $P_i \leftarrow \{a \in D_i \mid \forall j \in \partial a \setminus \{i\} : \hat{\boldsymbol{\sigma}}^{(1)}(i) = 0\}$
  $\hat{\boldsymbol{\sigma}}_{\texttt{SPOG}}(i) \leftarrow \mathbb{1}\left[|a \in P_i \cap \partial i \mid \tilde{\boldsymbol{\tau}}(a) = 1| \geq C \cdot |P_i|\right]$
return $\hat{\boldsymbol{\sigma}}_{\texttt{SPOG}}$

Algorithm 1: SPOG

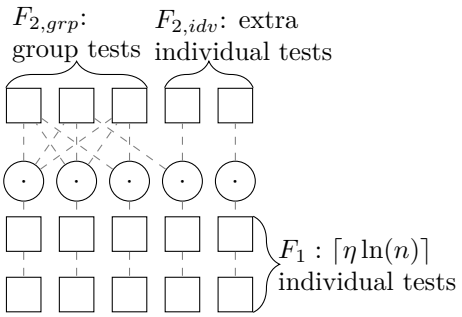


Figure 2: Illustration of $\boldsymbol{G}_{\texttt{SPOG}}$. Circles represent individuals and squares tests.

Lemma 18 *$\boldsymbol{G}_{\texttt{SPOG}}(\hat{\alpha}, \boldsymbol{p}, \Gamma, \eta, \varepsilon)$ uses at most $\left(\eta + \left(1 + \frac{\varepsilon}{3}\right)\frac{\xi(\alpha,\boldsymbol{p},\Gamma)}{\Gamma}\right) n \ln n + \mathcal{O}(n)$ tests w.h.p.*

We now turn to the analysis of SPOG's correctness. First, we see that the classification of the pseudo-genie $\hat{\boldsymbol{\sigma}}^{(1)}$ approximates $\boldsymbol{\sigma}$ well, in the sense that for each fixed individual i, it correctly identifies i's status (although of course, this does not apply for all individuals at the same time). The following lemma, proven in Appendix D.2.3 of the full version Hintze et al. (2024), collects this fact as well as some other useful properties of $\hat{\boldsymbol{\sigma}}^{(1)}$. We note that the value β stems from the choice of C as threshold.

Lemma 19 *For all $i \in [n]$: $\mathbb{P}\left[\hat{\boldsymbol{\sigma}}^{(1)}(i) \neq \boldsymbol{\sigma}(i)\right] \leq n^{-\beta\eta}$, and these events are independent. Furthermore, the upper bound on the probability also holds when conditioning on $\boldsymbol{\sigma}(i) = 1$ or $\boldsymbol{\sigma}(i) = 0$.*

Due to the restriction $\Gamma \leq \lceil \hat{\alpha}^{-1} \rceil$, it is just as unlikely that a pseudo-good test contains *any* individual misclassified by $\hat{\boldsymbol{\sigma}}^{(1)}$, i.e., that a pseudo-good test is *not* good. The proof of this bound in Appendix D.2.4 of the full version Hintze et al. (2024) additionally uses Lemma 19 as well as Bayes' theorem.

Lemma 20 *For any $i \in [n]$, the probability of a test $a \in P_i$ containing a misclassified individual (other than i) is in $\mathcal{O}(n^{-\beta\eta})$. To be precise: for all $i \in n$ and $a \in \partial i$,*

$$\mathbb{P}\left[\bigvee_{j \in \partial a \setminus \{i\}} \hat{\boldsymbol{\sigma}}^{(1)}(j) \neq \boldsymbol{\sigma}(j) \,\middle|\, a \in P_i\right] = \mathcal{O}(n^{-\beta\eta}).$$

This entails that the probability of an individual i being misclassified by SPOG is close to the probability of it being misclassified by the genie estimator if all pseudo-good tests were good. The proof of this lemma, found in Appendix A.2.5 of the full version Hintze et al. (2024), additionally uses the independence of observed test results P_i conditioned on $\sigma(i)$ (as they are distinctive by construction).

Lemma 21 *For any individual $i \in [n]$, $\mathbb{P}[\hat{\boldsymbol{\sigma}}_{\texttt{SPOG}}(i) \neq \boldsymbol{\sigma}(i) \mid |P_i|] \leq \exp\left(-|P_i|\left(\beta - \mathcal{O}(n^{-\beta\eta})\right)\right)$.*

To use this bound, we need a lower bound on the number of distinctive tests for each individual. The following bound is good enough for our purposes; its proof in Appendix D.2.6 of the full version Hintze et al. (2024) is rather similar to birthday-paradox-type calculations.

Lemma 22 *Let $\mathcal{D}$ be the event that for all $i \in [n]$, we have $|D_i| \geq (1 + \frac{\varepsilon}{6})\,\xi \ln n$. Then $\mathbb{P}[\mathcal{D}] = 1 - \mathcal{O}\left(\frac{\Gamma^4 \ln^4 n}{n}\right)$.*

The final lemma bounds the probability of misclassifying an individual (given that there are sufficiently many distinctive tests). In its proof (in Appendix D.2.7 of the full version Hintze et al. (2024)), we see that the probability of an individual i being misclassified is bounded by $\left(1 - (1-\alpha)^{\Gamma-1}(1 - e^{-\beta}) - \mathcal{O}(n^{-\delta})\right)^{|D_i|}$ given $|D_i|$. And ξ is defined to obtain the upper bound of the lemma as long as $|D_i| \geq (1 + \frac{\varepsilon}{6})\xi \ln n$.

Lemma 23 *Assume that $\Gamma = \mathcal{O}(n^{\beta\eta - \delta})$ for some $0 < \delta \leq \beta\eta$, and that $\alpha \leq \hat{\alpha}$. Then for $\mathcal{D}$ as in Lemma 22, we have, for all $i \in [n]$,*

$$\mathbb{P}[\hat{\boldsymbol{\sigma}}_{\mathit{SPOG}}(i) \neq \boldsymbol{\sigma}(i) \,|\, \mathcal{D}] = n^{-\left(1+\frac{\varepsilon}{6}\right)+\mathcal{O}(n^{-\delta})}.$$

Proof [Proof of Theorem 2] With these lemmas, the proof of Theorem 2 boils down to an appropriate choice of parameters. For details, see the proof of Lemma 4.17 in the full version of the paper Hintze et al. (2024). ■

5. Adaptive Group Testing

In the remainder of the allotted space, we give a brief overview of the main ideas used to prove Theorems 3 and 4. The full proof details are deferred to Section 5 of the full version of this paper Hintze et al. (2024).

Impossibility: proof strategy for Theorem 3 The high-level proof strategy is as follows: As in the proof of Theorem 1, we first modify the test scheme by adding a few individual tests to ease analysis, which can only increase the probability of successful estimation. We then define a notion of "typical" infection vectors: these are vectors for which, given a particular test design and displayed test results, most infected individuals have a number of good tests which is close to the expected number of good tests. First, we show that the ground truth $\boldsymbol{\sigma}$ is indeed "typical" w.h.p. for our modified test design. Then, we see that the posterior probability mass of a typical infection vector σ is dwarfed by that of a set of vectors that only differ in one coordinate from σ. This gives an upper bound on the posterior probability of σ. Combining all of the above then yields Theorem 3. The bound we obtain on the posterior odds ratio of a typical infection vector σ and neighbors of σ is similar in form to (Scarlett, 2019, inequality (37) in Lemma 2). However, our bound is stronger, enabling us to derive a sharper bound on the minimal number of tests required. This is why we obtain an additional factor of $(1 - 2p_{01})^{-1}$. To be precise, our approach improves the second element of the maximum in (Scarlett, 2019, inequality (34) in Lemma 2) by a factor of $(1 - 2p_{01})^{-1}$, even in sublinear cases.

The algorithm PRESTO: proof strategy for Theorem 4 The adaptive algorithm PRESTO runs three stages of tests. In the first stage, each individual is tested $\lceil \eta \ln(n) \rceil$ times individually (for a small η). The population is then partitioned via thresholding on the fraction of positive tests: the set S_1 contains almost only infected individuals, while the other, S_0, contains almost only uninfected individuals. In the second stage, all individuals in S_1 (of which there are $\approx \alpha n$) undergo further $\lceil (1 + \frac{\varepsilon}{4})\, m_{\mathrm{ad}}/(\alpha n) \rceil$ individual tests. Again, S_1 is partitioned via thresholding: the threshold is $p_{11} - \delta_1$ for a small δ_1 so that one set, U_1, contains *only* infected individuals w.h.p., while the other, U_0, again contains almost only healthy individuals. The third stage cleans up S_0 and U_0 by performing SPOG separately on both sets, using the variant for sub-constant α. We keep the sets separate to maintain an i.i.d. prior.

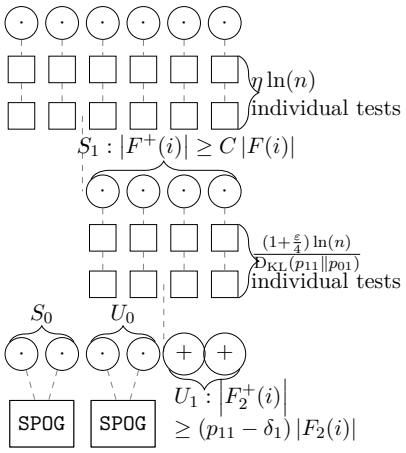


Figure 3: An illustration of the adaptive test scheme of PRESTO. The circles represent the individuals and the squares represent tests.

This multi-stage approach with clean-up steps is very similar to the three-stage algorithm for noisy group testing in Scarlett (2019). The number of tests each individual receives in our second stage is the same as that in Algorithm 2, Step 2*b* of Scarlett (2019). The proof of correctness of PRESTO consists of a combination of standard techniques as well as the correctness of SPOG in the sublinear regime.

Acknowledgments

Lena Krieg is supported by DFG CO 646/3 and DFG CO 646/5. Olga Scheftelowitsch is supported by DFG CO 646/5. Haodong Zhu is supported by the NWO Gravitation project NETWORKS under grant no. 024.002.003 and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 945045. We thank Amin Coja-Oghlan and Mihyun Kang for organizing and inviting us to a workshop in Strobl in 2023, where we initiated the collaboration that led to this work. We also thank Jonathan Scarlett for a valuable comment on a previous version of this work.

References

Emmanuel Abbe. Community detection and stochastic block models: Recent developments. *J. Mach. Learn. Res.*, 18:177:1–177:86, 2017. URL https://jmlr.org/papers/v18/16-480.html.

Emmanuel Abbe and Colin Sandon. Community detection in general stochastic block models: fundamental limits and efficient recovery algorithms. *CoRR*, abs/1503.00609, 2015. URL http://arxiv.org/abs/1503.00609.

Matthew Aldridge. Individual testing is optimal for nonadaptive group testing in the linear regime. *IEEE Trans. Inf. Theory*, 65(4):2058–2061, 2019a.

Matthew Aldridge. Rates of adaptive group testing in the linear regime. *IEEE International Symposium on Information Theory (ISIT)*, pages 236–240, 2019b.

Matthew Aldridge. Conservative two-stage group testing in the linear regime. *arXiv preprint arXiv:2005.06617*, 2020.

Matthew Aldridge, Leonardo Baldassini, and Oliver Johnson. Group testing algorithms: Bounds and simulations. *IEEE Transactions on Information Theory*, 60(6):3671–3687, 2014. doi:10.1109/TIT.2014.2314472.

Matthew Aldridge, Oliver Johnson, and Jonathan Scarlett. Group testing: An information theory perspective. *Found. Trends Commun. Inf. Theory*, 15(3-4):196–392, 2019. doi:10.1561/0100000099. URL https://doi.org/10.1561/0100000099.

HB Chen, Ding-Zhu Du, and Frank K Hwang. An unexpected meeting of four seemingly unrelated problems: graph testing, dna complex screening, superimposed codes and secure key distribution. *Journal of Combinatorial Optimization*, 14(2):121–129, 2007.

Junren Chen and Jonathan Scarlett. Exact thresholds for noisy non-adaptive group testing. *arXiv:2401.04884*, 2024. To appear in: Proc. of 36th ACM-SIAM Symposium on Discrete Algorithms, SODA (2025).

Mahdi Cheraghchi and Vasileios Nakos. Combinatorial group testing and sparse recovery schemes with near-optimal decoding time. *proc. 61st IEEE Annual Symposium on Foundations of Computer Science, FOCS*, pages 1203–1213, 2020.

Amin Coja-Oghlan, Oliver Gebhard, Max Hahn-Klimroth, and Philipp Loick. Information-theoretic and algorithmic thresholds for group testing. *IEEE Transactions on Information Theory*, 66(12):7911–7928, 2020.

Amin Coja-Oghlan, Oliver Gebhard, Max Hahn-Klimroth, and Philipp Loick. Optimal group testing. *Comb. Probab. Comput.*, 30(6):811–848, 2021. doi:10.1017/S096354832100002X. URL https://doi.org/10.1017/S096354832100002X.

Amin Coja-Oghlan, Max Hahn-Klimroth, Lukas Hintze, Dominik Kaaser, Lena Krieg, Maurice Rolvien, and Olga Scheftelowitsch. Noisy group testing via spatial coupling. *arXiv:2402.02895*, 2024.

Robert Dorfman. The Detection of Defective Members of Large Populations. *The Annals of Mathematical Statistics*, 14(4):436 – 440, 1943. doi:10.1214/aoms/1177731363. URL https://doi.org/10.1214/aoms/1177731363.

Uriel Feige, Prabhakar Raghavan, David Peleg, and Eli Upfal. Computing with noisy information. *SIAM Journal on Computing*, 23(5):1001–1018, 1994. doi:10.1137/S0097539791195877. URL https://doi.org/10.1137/S0097539791195877.

Karin Haar, Thomas Meyer, Sarika Desai, Michael Thamm, Matthias an der Heiden, Viviane Bremer, and Osamah Hamouda. Low sensitivity of pooled chlamydia testing in a sample of the young german general population. *Journal of US-China Medical Science*, 8(10), 2011.

Lukas Hintze, Lena Krieg, Olga Scheftelowitsch, and Haodong Zhu. Noisy linear group testing: Exact thresholds and efficient algorithms. *arXiv:2411.03839*, 2024. URL https://arxiv.org/abs/2411.03839.

Edwin S Hong, Richard E Ladner, and Eve A Riskin. Group testing for wavelet packet image compression. *Proc. IEEE Data Compression Conference DCC*, pages 73–82, 2001.

F. K. Hwang. A method for detecting all defective members in a population by group testing. *Journal of the American Statistical Association*, 67(339):605–608, 1972. doi:10.1080/01621459.1972.10481257. URL https://www.tandfonline.com/doi/abs/10.1080/01621459.1972.10481257.

F. K. Jasima, Nilofer Tutor, and P. Subhashini. Comparative evaluation of elisa & clia screening assays in the effective detection of hiv infection in blood donor samples. *International journal of health sciences*, 2022. URL https://api.semanticscholar.org/CorpusID:249326416.

Oliver Johnson, Matthew Aldridge, and Jonathan Scarlett. Performance of group testing algorithms with near-constant tests per item. *IEEE Transactions on Information Theory*, 65(2):707–723, 2018.

Elmas Pınar Kahraman Kılbaş, Imdat Kilbas, and Ihsan Hakki Ciftci. A meta-analysis on the comparison of the sensitivity of three test methods used in the diagnosis of covid-19. *Journal of Kermanshah University of Medical Sciences*, 2022. URL https://api.semanticscholar.org/CorpusID:253491551.

Leon Mutesa, Pacifique Ndishimye, Yvan Butera, Jacob Souopgui, Annette Uwineza, Robert Rutayisire, Ella Larissa Ndoricimpaye, Emile Musoni, Nadine Rujeni, Thierry Nyatanyi, et al. A pooled testing strategy for identifying sars-cov-2 at low prevalence. *Nature*, 589 (7841):276–280, 2021.

Jonathan Scarlett. Noisy adaptive group testing: Bounds and algorithms. *IEEE Trans. Inf. Theory*, 65(6):3646–3661, 2019. doi:10.1109/TIT.2018.2883604. URL https://doi.org/10.1109/TIT.2018.2883604.

Jonathan Scarlett and Volkan Cevher. Phase transitions in group testing. *Proc. 27th ACM-SIAM symposium on Discrete algorithms*, pages 40–53, 2016.

Anirudh K. Singh, Ram Kumar Nema, Ankur Joshi, Prem Shankar, Shashwati Nema, Arun K Raghuwanshi, Chitra Patankar, Bijina J. Mathew, Arti Shrivas, Ritu Pandey, Ranu Tripathi, Debasis Biswas, and Sarman Singh. Evaluation of pooled sample analysis strategy in expediting case detection in areas with emerging outbreaks of covid-19: A pilot study. *PLoS ONE*, 15, 2020. URL https://api.semanticscholar.org/CorpusID:221862777.

Milton Sobel and Phyllis A Groll. Group testing to eliminate efficiently all defectives in a binomial sample. *Bell System Technical Journal*, 38(5):1179–1252, 1959.

Tuna Toptan, Lisa Eckermann, Annika E Pfeiffer, Sebastian Hoehl, Sandra Ciesek, Christian Drosten, and Victor M Corman. Evaluation of a sars-cov-2 rapid antigen test: Potential to help reduce community spread? *Journal of Clinical Virology*, 135:104713, 2021.

Peter Ungar. The cutoff point for group testing. *Communications on Pure and Applied Mathematics*, 13:49–54, 1960.

Jack Wolf. Born again group testing: Multiaccess communications. *IEEE Transactions on Information Theory*, 31(2):185–191, 1985.