View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Explore Bristol Research

Scarlett, J., & Johnson, O. T. (2020). Noisy Non-Adaptive Group Testing: A (Near-)Definite Defectives Approach. *IEEE Transactions on Information Theory*, *66*(6), 3775-3797. https://doi.org/10.1109/TIT.2020.2970184

Peer reviewed version

Link to published version (if available):
10.1109/TIT.2020.2970184

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via IEEE at https://ieeexplore.ieee.org/document/8974412. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

Noisy Non-Adaptive Group Testing: A (Near-)Definite Defectives Approach

Jonathan Scarlett and Oliver Johnson

***Abstract*—The group testing problem consists of determining a small set of defective items from a larger set of items based on a number of possibly-noisy tests, and is relevant in applications such as medical testing, communication protocols, pattern matching, and more. We study the noisy version of this problem, where the outcome of each standard noiseless group test is subject to independent noise, corresponding to passing the noiseless result through a binary channel. We introduce a class of algorithms that we refer to as Near-Definite Defectives (NDD), and study bounds on the required number of tests for asymptotically vanishing error probability under Bernoulli random test designs. In addition, we study algorithm-independent converse results, giving lower bounds on the required number of tests under Bernoulli test designs. Under reverse Z-channel noise, the achievable rates and converse results match in a broad range of sparsity regimes, and under Z-channel noise, the two match in a narrower range of dense/low-noise regimes. We observe that although these two channels have the same Shannon capacity when viewed as a communication channel, they can behave quite differently when it comes to group testing. Finally, we extend our analysis of these noise models to a general binary noise model (including symmetric noise), and show improvements over known existing bounds in broad scaling regimes.**

***Index Terms*—Group testing, performance bounds, sparsity, Z channel, information-theoretic limits**

I. Introduction

The group testing problem consists of determining a small subset of "defective" items within a larger set of items, based on a number of possibly-noisy tests. As described in more detail in the survey monograph [1], this problem has a history in medical testing [2], and has regained significant attention with subsequent applications in areas such as communication protocols [3], DNA sequencing [4], data forensics [5], pattern matching [6], and database systems [7], as well as new connections with compressive sensing [8], [9]. The general setup involves a sequence of tests, each of which acts on a particular subset (or "pool") of items and produces an outcome Y that can be a deterministic or random function of the defectivity status of the items in the pool.

Jonathan Scarlett is with the Department of Computer Science, National University of Singapore, Singapore 117417, and also with the Department of Mathematics, National University of Singapore, Singapore 119076 (e-mail: scarlett@comp.nus.edu.sg).

Oliver Johnson is with the School of Mathematics, University of Bristol, UK (e-mail: maotj@bristol.ac.uk).

In recent years, the information-theoretic limits and performance limits of practical algorithms for *noiseless* group testing have become increasingly well-understood [10]–[16]. By comparison, random noise settings are somewhat less well-understood despite ongoing advances [13], [17]–[20]. In particular, the algorithm that gives the best known *noiseless* performance guarantees in most sparsity regimes (among practical algorithms), known as Definite Defectives (DD) [12], [15], has no previous noisy counterpart. In this paper, we address this gap by introducing and studying noisy variants of DD, and showing that they provide the best known performance bounds in a wide range of settings depending on the sparsity and noise level.

A. Overview of Noiseless Group Testing

Let p denote the number of items, and let $S \subseteq \{1,\dots,p\}$ denote the set of defective items. In the standard noiseless setting first introduced in [2], the outcome of each test takes the form

$$Y = \bigvee_{j \in S} X_j, \tag{1}$$

where the test vector $X = (X_1,\dots,X_p) \in \{0,1\}^p$ indicates which items are included in the test. That is, the resulting outcome $Y = 1$ if and only if at least one defective item was included in the test. We refer to tests with $Y = 1$ as positive, and tests with $Y = 0$ as negative.

Given the tests and their outcomes, a *decoder* forms an estimate $\widehat{S}$ of S. One wishes to design a sequence of tests $X^{(1)},\dots,X^{(n)}$, with n ideally as small as possible, such that the decoder recovers S with probability arbitrarily close to one. The error probability is given by

$$P_{\mathrm{e}} := \mathbb{P}[\widehat{S} \neq S], \tag{2}$$

and is taken over the randomness of the defective set S, the tests $X^{(1)},\dots,X^{(n)}$ (if randomized), and the test outcomes $Y^{(1)},\dots,Y^{(n)}$ (if noisy). For convenience, we represent the tests as a matrix $\mathbf{X} \in \{0,1\}^{n\times p}$, where the i-th row is $X^{(i)}$ and represents the i-th test.

In this paper, we consider the case that, for a given sparsity level k, the defective set S is chosen uniformly on the $\binom{p}{k}$ subsets of $\{1,\dots,p\}$ of cardinality k. Following

recent works such as [11], we define the rate (in bits/test) of a group-testing algorithm using n tests to be[1]

$$R := \frac{\log_2 \binom{p}{k}}{n}, \tag{3}$$

which we can think of as the number of bits of information about the defective set learned per test. We consider the asymptotic regime where $p \to \infty$ with $k \asymp p^\theta$ for some $\theta \in (0,1)$,[2] so we will often use the equivalent limiting definition that

$$R \sim \frac{k \log_2(p/k)}{n}, \tag{4}$$

where $\sim$ denotes asymptotic equality up to a multiplicative $1 + o(1)$ term.

It is well known from standard information-theoretic arguments (e.g., [11]) that no algorithm with rate above 1 bit/test can have vanishing error probability. For noiseless adaptive group testing (where the choice of test $X^{(i+1)}$ can depend on the previous tests $X^{(1)}, \dots, X^{(i)}$ and their outcomes $Y^{(1)}, \dots, Y^{(i)}$), Hwang's algorithm [21] has error probability tending to zero with rate approaching one in any regime where $k = o(p)$, and is therefore asymptotically optimal.

In this paper, we study non-adaptive group testing, where the entire collection of tests $\mathbf{X} \in \{0,1\}^{n \times p}$ is fixed in advance. We focus in particular on *Bernoulli testing*, where the entries of $\mathbf{X}$ are independently drawn at random from $\mathrm{Bernoulli}\big(\frac{\nu}{k}\big)$ for some parameter $\nu > 0$.

In the noiseless case, it is known that the non-adaptive definite defectives (DD) algorithm [22] is both practically implementable (in terms of storage and processing requirements) and performs well in terms of rate. Specifically, under Bernoulli testing in the regime $k \asymp p^\theta$, the DD algorithm achieves $P_{\mathrm{e}} \to 0$ when [12, Theorem 12]

$$R < \frac{1}{e \log 2} \min\left\{1, \frac{1-\theta}{\theta}\right\}. \tag{5}$$

Furthermore, the DD algorithm is known to be rate-optimal for sufficiently dense problems (specifically, for $\theta > 1/2$) under Bernoulli testing, in the sense that *any* algorithm has P_{e} bounded away from zero if the rate satisfies [12], [14]

$$R > \frac{1}{e \log 2} \frac{1-\theta}{\theta}. \tag{6}$$

The contribution of this paper is to extend the bounds of the form (5) and (6) to noisy group testing models, by introducing and referring to a class of algorithms which we refer to as noisy DD (NDD). For clarity, we first review both (noiseless) DD [22] and a related algorithm called COMP [23], which forms the first stage of DD.

[1]Throughout the paper, log refers to natural logarithms taken to base e, and we write $\log_2$ for base 2 logarithms.

[2]Here and subsequently, $k \asymp p^\theta$ means that $\frac{k}{p^\theta}$ is bounded away from both 0 and ∞ in the limit as $p \to \infty$.

Definition 1. The Combinatorial Orthogonal Matching Pursuit (COMP) and Definite Defectives (DD) algorithms for *noiseless* non-adaptive group testing are defined as follows:

1. Since $Y = 1$ if and only if the test pool contains a defective item, we can be sure that each item that appears in a negative test is not defective. We can form a list of such items formed from all tests, which we refer to as ND; the rest of the items $\mathrm{PD} := \{1, \dots, p\} \backslash \mathrm{ND}$ are considered "possibly defective". The COMP algorithm simply estimates S using the set of possible defective items, $\widehat{S} = \mathrm{PD}$.
2. The DD algorithm starts with the possible defective items PD. Since every positive test must contain at least one defective item, if a test with $Y = 1$ contains exactly one item from PD, then we can be certain that the item in question is defective. The DD algorithm outputs $\widehat{S}$ equaling the set of PD items that appear in a positive test with no other PD item.

B. Noisy Group Testing

Generalizing (1), we consider noisy models that correspond to passing the quantity $U = \vee_{j \in S} X_j$ through a noisy channel $P_{Y|U}$. We focus in particular on the following special cases, each of which depends on one or two noise parameters that can be set to zero to recover the noiseless model.

Definition 2. We define the following noise models, illustrated in Figure 1:

1. The general binary channel model is given by
$$\begin{aligned} P_{Y|U}(0|0) &= 1 - \rho_{01}, \quad P_{Y|U}(1|0) = \rho_{01}, \\ P_{Y|U}(0|1) &= \rho_{10}, \quad P_{Y|U}(1|1) = 1 - \rho_{10} \end{aligned} \tag{7}$$
for some noise levels ρ_{01} and ρ_{10} both in $[0,1]$.
2. The Z-channel model is obtained by taking $\rho_{01} = 0$ and $\rho_{10} = \rho$ in the general model, yielding so that
$$\begin{aligned} P_{Y|U}(0|0) &= 1, \quad P_{Y|U}(1|0) = 0, \\ P_{Y|U}(0|1) &= \rho, \quad P_{Y|U}(1|1) = 1 - \rho \end{aligned} \tag{8}$$
for some noise level $\rho \in [0,1]$. Hence, if $Y = 1$, then the test must contain a defective item.
3. The reverse Z-channel (RZ-channel) model, also known as the addition noise model [10], is obtained by taking $\rho_{01} = \rho$ and $\rho_{10} = 0$, yielding
$$\begin{aligned} P_{Y|U}(0|0) &= 1 - \rho, \quad P_{Y|U}(1|0) = \rho, \\ P_{Y|U}(0|1) &= 0, \quad P_{Y|U}(1|1) = 1 \end{aligned} \tag{9}$$
for some noise level $\rho \in [0,1]$. Hence, if $Y = 0$, then the test must contain no defective items.
4. The symmetric noise model is obtained by taking $\rho_{01} = \rho_{10} = \rho$ in the general model, yielding
$$P_{Y|U}(y|u) = \begin{cases} 1 - \rho & y = u \\ \rho & y \neq u \end{cases} \tag{10}$$

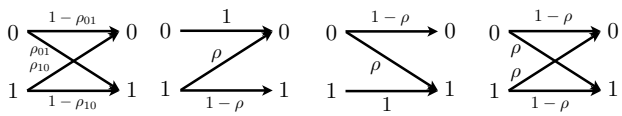

Figure 1: General binary channel, Z-channel, reverse Z-channel, and binary symmetric channel.

for some noise level $\rho \in [0, 1]$.

In this paper, we will focus primarily on the Z and reverse Z-channel models, applications of which are discussed in Section I-D. As well as being important in their own right for applications, these models serve as useful stepping stones towards the general binary model (including the symmetric model), which is handled in Appendix I.

Note that for the general binary channel model, we can assume without loss of generality that $\rho_{01}+\rho_{10} \le 1$. This is because (as for a standard binary symmetric channel) if this were not true, we could flip the outcome of all tests as a pre-processing step, producing a new channel with $\widetilde{\rho}_{01} = 1-\rho_{01}$ and $\widetilde{\rho}_{10} = 1-\rho_{10}$ satisfying $\widetilde{\rho}_{01}+\widetilde{\rho}_{01} \le 1$. Additionally, observe that the case $\rho_{01} + \rho_{10} = 1$ is a degenerate one, under which the conditional distribution $P_{Y|U}(y|u)$ does not depend on u, meaning that U and Y are independent, so we cannot hope to recover useful information about the defective set from Y.

When considered to define a standard noisy communication channel, both the Z-channel and reverse Z-channel have Shannon capacity (in bits/use) given by [24]

$$C_{\mathrm{Z}}(\rho) = \log_2\left(1 + (1-\rho)\rho^{\rho/(1-\rho)}\right), \tag{11}$$

and the symmetric noise model has Shannon capacity (in bits/use) given by

$$C_{\mathrm{BSC}}(\rho) = 1 - h(\rho) \tag{12}$$

where $h(\rho) = -\rho\log_2\rho - (1-\rho)\log_2(1-\rho)$ is the binary entropy in bits. Since the Z-channel and reverse Z-channel have the same Shannon capacity, the information-theoretic results of [13], [25] suggest that they may require the same asymptotic number of tests, at least for sufficiently sparse settings. On the other hand, when adopting an NDD approach, it is unclear *a priori* which model requires more tests. See Section II-E for further discussion.

Except where stated otherwise, we assume that the noise levels ρ_{10}, ρ_{01} and number of defectives k are known; our analysis can also be applied to cases where only bounds are known, but the details become more tedious. Our main goal is to provide explicit achievable rates and converse bounds for noisy group testing under Bernoulli designs.

Remark 1. While the general binary model (7) captures several symmetric and non-symmetric noise models, there are other noise models of interest that it does not capture. As discussed in [1, Sec. 3.1], some noise models of interest depend on the *number of defectives* in the test, and not just the presence vs. absence of any defectives. A prominent example is dilution noise [10], in which each defective item in the test is independently "diluted" (and hence behaves as though it were non-defective) with some probability u. Hence, if there are ℓ defectives in the test, the probability they are all diluted is u^ℓ. While it may be possible to handle noise models of this kind using our techniques, the analysis appears to become significantly more complicated. This is primarily due to the different conditional distribution of Y for all different values of the number of defectives in the test, $\ell \in \{0, 1, \ldots, k\}$, in contrast with (7) in which only need to distinguish between $\ell = 0$ and $\ell \ge 1$. Due to these complications, we leave these further generalizations for future work.

C. Related Work

The information-theoretic limits of noiseless and noisy non-adaptive group testing were initially studied in the Russian literature [25], [26], and have recently become increasingly well-understood [10], [13], [14], [22], [27], [28]. Among the existing works, the results most relevant to the present paper are as follows:

- For both the adaptive and non-adaptive settings, it was shown by Baldassini *et al.* [11] that if the outcome Y is produced by passing the noiseless outcome $U = \vee_{j\in S} X_j$ through a channel $P_{Y|U}$, then any group testing achieving $P_{\mathrm{e}} \to 0$ the must have rate $R \le C$, where R is defined in (3) and C is the Shannon capacity of $P_{Y|U}$. Equivalently, the number of tests must satisfy $n \ge \left(\frac{1}{C} k \log_2 \frac{p}{k}\right)(1-o(1))$. For instance, under the symmetric noise model (10), this yields
$$n \ge \frac{k\log_2\frac{p}{k}}{1-h(\rho)}(1-o(1)). \tag{13}$$
It has recently been shown that under the RZ and symmetric noise models, this converse is not tight (i.e., it can be improved) when $\theta \in (0, 1)$ is sufficiently close to one, even in the adaptive setting [20].
- In the non-adaptive setting with symmetric noise, it was shown in [13], [27] that an information-theoretic threshold decoder attains the bound (13) when $k \asymp p^\theta$ for *sufficiently small* $\theta > 0$. The analysis of [20, Appendix A] shows that analogous findings also hold for the Z and RZ noise models.

Several non-adaptive noisy group testing algorithms have been shown to come with rigorous guarantees.

- The *Noisy Combinatorial Orthogonal Matching Pursuit* (NCOMP) algorithm checks, for each item, the proportion of tests it was included in that returned

positive, and declares the item to be defective if this number exceeds a suitably-chosen threshold. This is known to provide optimal scaling laws for the regime $k \asymp p^\theta$ ($\theta \in (0,1)$) [17], [23], albeit with somewhat suboptimal constants. That is, in the terminology of (3), NCOMP has a non-zero but suboptimal rate under symmetric noise. Similar results are also obtained for the general binary noise channel using a linear programming based algorithm in [17, Thm. 7].
- The method of *separate decoding of items*, also known as *separate testing of inputs* [19], [25], also considers the items separately, but uses all of the tests. Specifically, a given item's status is selected via a binary hypothesis test. This method was studied for $k = O(1)$ in [25], and for $k \asymp p^\theta$ in [19]. In particular, it was shown that for the symmetric noise model, the number of tests is within a factor $\log 2$ of the optimal information-theoretic threshold as $\theta \to 0$. However, the rate quickly become weaker as θ increases away from zero; see Appendix I for an example.

Since other works on noisy group testing are less related to the present paper, we only provide a brief outline. Some heuristic algorithms have been proposed for noisy settings without theoretical guarantees, including belief propagation [29] and a noisy linear programming relaxation [30]. Sublinear-time algorithms with guarantees on the number of samples and runtime have been proposed [31]–[34] (see also the earlier works of [35]–[37]), but the constants (and sometimes logarithmic factors) in the sample complexity bounds are far from optimal. The complementary viewpoint of *adversarial noise* has also been explored [35], [36], [38].

D. Applications of Noise Models

Group testing has been applied in a wide range of contexts, including biology, communications, information technology and data science, as outlined in [1, Section 1.7]. While the noiseless model has been widely studied from a theoretical point of view, in many of these applications it is unrealistic to assume that all tests will return a perfectly accurate answer. Many papers have dealt with this issue by studying the symmetric noise model described in (10).

However, we believe that in many applications, this assumption of symmetry is itself also unrealistic. Since there are often different mechanisms operating that may "flip" positive tests to negative and vice versa, there is no *a priori* reason to believe that these two types of error should be equally likely. This argument motivates the general binary channel model of Definition 2. In addition, to motivate the study of the RZ and Z channels in their own right, we proceed by giving examples where these models naturally arise.

In the *file comparison* problem, we wish to carry out data forensics to determine which out of a collection of computer files have been changed. One way to do this, described in [5], is to store a number of hashes of various concatenated collections of files. By comparing the hashes before and after any possible tampering, if the hash has changed, we know that at least one file in the group has been altered. This can be thought of as a group testing scenario: An altered file corresponds to a defective item, the collection of files corresponds to the testing pool, and a changed hash corresponds to a positive test. However, as discussed in [39], it is possible that the files may be altered in a way that does not change the value of the hash. In this sense, a test that should be positive may fail to be detected as such with a certain probability – this is exactly the Z-channel of Definition 2. While [39] argues that this effect can be minimized by taking arbitrarily long hashes, it may be that from an efficiency point of view it is preferable to store shorter hashes and take into account the effect of the Z channel noise in identifying the modified files.

An application for which the RZ-channel can serve as a natural model is that of *multiple-access communication*, e.g., see [3], [40], [41]. Roughly speaking, group testing permits the detection of a small subset of active users by requesting each user to transmit a signal at the times corresponding to 1's in the group testing matrix. Then, obtaining the group testing outcomes only requires detecting whether or not there is any signal present at each time instant. If this detection procedure is reliable, then there should be no false negatives; however, in the presence of an *interfering signal*, one is prone to false positives, in agreement with the RZ channel model of Definition 2. Alternatively, if the detection procedure is not perfectly reliable, then we may be subject to the general binary noise model with suitably-chosen values of ρ_{01} and ρ_{10}.

II. Summary of Main Results

As mentioned above, we will provide achievable rates for noisy group testing using noisy variants of the COMP and DD algorithms, as well as providing algorithm-independent converse bounds. We summarize our main results in the following subsections.

A. Highlights

Since the statements of our main results are somewhat technical, we begin by highlighting some key special cases, focusing primarily on scenarios in which our bounds are tight in a certain sense. The relevant rates are plotted in Figures 3 and 5 for the RZ- and Z-channel models respectively, and in Figure 7 in Appendix I for the symmetric noise model. We have the following:

- For the RZ-channel model, in Theorem 1 we establish an achievable rate and an algorithm-independent

converse (for Bernoulli testing) that match in broad scaling regimes, e.g., for all $\theta > 0.212$ in the case that $\rho = 0.1$. As the noise level ρ increases, the bounds match for a broader range of θ.

- Also for the RZ-channel model, in Appendix H we show that for θ close to zero and ρ close to one (i.e., the noisy and sparse setting), the achievable rate is approximately $\frac{1-\rho}{e\log 2}$; according to a simple capacity-based converse [11], this cannot be improved by any *arbitrary and possibly adaptive* algorithm.
- For the Z-channel model, in Theorem 2 we establish an achievable rate and an algorithm-independent converse (for Bernoulli testing) that match in certain dense scaling regimes when ρ is small enough, e.g., for all $\theta > 0.729$ in the case that $\rho = 0.001$. However, as the noise level ρ increases, the bounds in fact match for a narrower range of θ (and eventually for no θ).
- For each noise model, we show that we recover the rate of [12] for the noiseless setting in the limit as the noise parameters tend to zero; this rate is known to be tight for Bernoulli testing whenever $\theta > \frac{1}{2}$.

B. Preliminaries

Notation. First, we establish some additional notation.

Definition 3. For any $\gamma > 0$, we define the function

$$D_\gamma(t) = t\log\left(\frac{t}{\gamma}\right) - t + \gamma \qquad \text{for } t \geq 0. \tag{14}$$

Note that $D_\gamma(\gamma) = D'_\gamma(\gamma) = 0$, and that $D''_\gamma(t) = 1/t \geq 0$, so D_γ is convex. Hence, $D_\gamma(t) \geq 0$ for all $t \geq 0$, and $D_\gamma(t)$ is strictly increasing for $t > \gamma$. Note further that

$$D_a(t) = aD_1(t/a) \qquad \text{for all } a > 0 \text{ and } t \geq 0. \tag{15}$$

At several points in the paper, we require explicit values for the intersection of two functions related to the $D_\gamma(t)$ function (for different choices of γ). This intersection is found in Lemma 7 of Appendix A, and can be expressed in terms of the Lambert W-function (see for example [42]). This function gives the solution to the equation $W(x)e^{W(x)} = x$ for $x \geq -1/e$, and has two real branches at the point $(-1/e, -1)$; see Figure 2. We shall write W_0 and W_{-1} respectively for the principal branch ($W_0(x) \geq -1$) and lower branch ($W_{-1}(x) \leq -1$). The key properties of W that we shall require are the derivative [42, Eq. (3.2)]

$$W'(x) = \frac{W(x)}{x(1+W(x))} \qquad \text{for } x \notin \{0, -1/e\} \tag{16}$$

(which holds on either branch), and the following asymptotic expansions, [42, Section 4]:

$$W_0(x) = \log x - \log\log x + o(1), \quad x \to \infty \tag{17}$$

$$W_{-1}(x) = \log(-x) - \log(-\log(-x)) + o(1), \quad x \to 0^-, \tag{18}$$

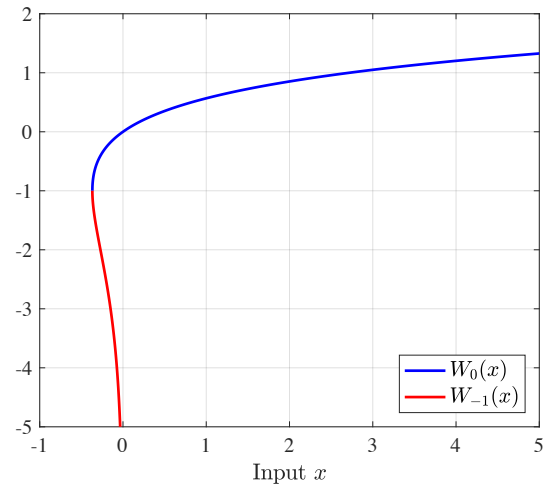

Figure 2: Upper and lower branches of the Lambert W-function.

where $x \to 0^-$ means approaching zero from below. Some intuition behind these expansion is as follows: Direct calculation using the fact that $W(x)e^{W(x)} = x$ means we can deduce that when $W(x) \geq 1$, we have $x \geq e^{W(x)}$, and hence $W_0(x) \leq \log x$ for $x \geq e$. Similarly, when $W(x) \leq -1$, we have $e^{W(x)} \leq -x$, and hence $W_{-1}(x) \leq \log(-x)$ for all x.

Concentration Bounds. We consider D_γ of Definition 3 because it naturally arises in tail bounds on binomial random variables. Specifically, we will use the following [43, Ch. 4]: For $Z \sim \text{Binomial}(N, q)$, we have that for any $\epsilon > 0$ that

$$\mathbb{P}[Z \leq Nq(1-\epsilon)] \leq \exp\Big(-NqD_1(1-\epsilon)\Big), \tag{19}$$

$$\mathbb{P}[Z \geq Nq(1+\epsilon)] \leq \exp\Big(-NqD_1(1+\epsilon)\Big). \tag{20}$$

The bounds of (19) and (20) are asymptotically tight in certain regimes. To establish this fact, we will make use of the following binomial coefficient bound [44, Lemma 4.7.1]:

$$\binom{N}{\delta N} \geq \frac{1}{\sqrt{8N\delta(1-\delta)}}\exp(Nh_e(\delta)),$$

where h_e is the binary entropy function in nats. Using this bound, we deduce that for given $\phi \in (0,1)$ such that

$Nq\phi$ is an integer, we have

$$\begin{aligned}
&\mathbb{P}[Z = Nq\phi] \\
&= \binom{N}{Nq\phi} q^{Nq\phi}(1-q)^{N(1-q\phi)} && (21)\\
&\geq \frac{1}{\sqrt{8Nq\phi(1-q\phi)}} \exp\Bigg(-NqD_1(\phi) \\
&\quad + N\Bigg(q(1-\phi) + (1-q\phi)\log\Big(\frac{1-q}{1-q\phi}\Big)\Bigg)\Bigg) && (22)\\
&\geq \frac{\exp\big(-N(1-\phi)^2q^2/(1-q)\big)}{\sqrt{8Nq\phi(1-q\phi)}} \exp\big(-NqD_1(\phi)\big), && (23)
\end{aligned}$$

where (22) uses $qD_1(\phi) = q\phi\log\phi + q(1-\phi)$, and (23) follows from substituting the value $u = \frac{1-q}{1-q\phi}$ in the bound $\log u \geq 1 - \frac{1}{u}$ and rearranging.

We refer to the term preceding the exponential in (23) as the "sharpness factor". Observe that if ϕ is constant and we have $Nq = \big(a\log p\big)(1+o(1))$, $q = \frac{b}{k}(1+o(1))$, and $k \asymp p^\theta$ for some positive constants a, b, θ, then for any $\epsilon' > 0$ this sharpness factor is lower bounded by $p^{-\epsilon'}$ for p sufficiently large. By picking $\phi = 1 \pm \epsilon$, and bounding the tail by the respective point probability, we deduce that (19) and (20) are each tight to within this sharpness factor.

C. Reverse Z-Channel

Our main result for the reverse Z-channel is written in terms of the following technical definitions. First, we define

$$\kappa = \kappa(\theta) := -W_{-1}\left(-e^{-1}\rho^{\theta/(1-\theta)}\right), \tag{24}$$

where W_{-1} denotes the lower branch of the Lambert W-function; that is, $\kappa > 0$ is a solution to the equation $\kappa e^{-\kappa} = e^{-1}\rho^{\theta/(1-\theta)}$. Moreover, we write

$$\theta_{\text{crit}}^{(\text{RZ})} := \theta_{\text{crit}}^{(\text{RZ})}(\rho) = 1 + \frac{\rho\log\rho}{1-\rho} \tag{25}$$

$$\theta_{\text{opt}} := \theta_{\text{opt}}(\rho) = \frac{t(\rho)}{\log\rho + t(\rho)}, \tag{26}$$

where $t(\rho) := -\log(1-\rho) + \log(-\log(\rho)) + \frac{\log(\rho)}{1-\rho} + 1$.

Theorem 1. (Reverse Z-Channel) *For noisy group testing under reverse Z-channel noise with parameter $\rho \in (0,1)$, in the regime where $k \asymp p^\theta$ with $\theta \in (0,1)$, we have the following under Bernoulli testing:*

1. **[Achievability]** *Under the Bernoulli testing parameter $\nu = 1$, there exists a practical algorithm achieving error probability $P_{\text{e}} \to 0$ with rate*

$$\underline{R}^{(\text{RZ})}(\theta,\rho) = \begin{cases} \frac{1-\rho}{e\log 2}, & \theta \leq \theta_{\text{opt}}, \\ \frac{-\log\rho}{\kappa(\theta)e\log 2}, & \theta_{\text{opt}} \leq \theta \leq \theta_{\text{crit}}^{(\text{RZ})}, \\ \frac{(1-\theta)(1-\rho)}{e\log 2}, & \theta \geq \theta_{\text{crit}}^{(\text{RZ})}. \end{cases} \tag{27}$$

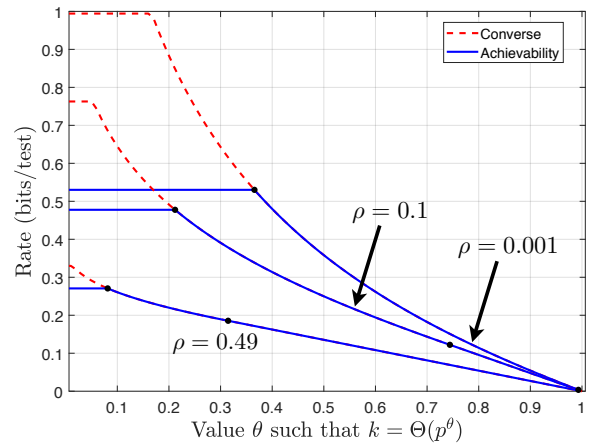

Figure 3: Achievable rate and algorithm-independent lower bound under reverse Z-channel noise and Bernoulli testing. The dots indicate the thresholds θ_{opt} and $\theta_{\text{crit}}^{(\text{RZ})}$.

2. **[Converse]** *Under any Bernoulli testing parameter $\nu > 0$, if $\rho < 1/2$, then no algorithm can achieve $P_{\text{e}} \to 0$ with a rate higher than*

$$\overline{R}^{(\text{RZ})}(\theta,\rho) = \begin{cases} \min\Big\{C_Z(\rho), \frac{-\log\rho}{\kappa(\theta)e\log 2}\Big\}, & \theta \leq \theta_{\text{crit}}^{(\text{RZ})}, \\ \frac{(1-\theta)(1-\rho)}{e\log 2}, & \theta \geq \theta_{\text{crit}}^{(\text{RZ})}. \end{cases} \tag{28}$$

where $C_Z(\rho)$ is the Shannon capacity of the channel, given by (11).

Proof. See Section III. □

These rates are illustrated for three different noise levels in Figure 3.

Remark 2.

1. As in the noiseless case, for sufficiently dense problems (i.e., for $\theta \geq \theta_{\text{opt}}$), we obtain a sharp result, with the achievable and converse rates coinciding. In this case, this optimal performance is achieved by either a noisy version of the DD algorithm (for $\theta_{\text{opt}} \leq \theta \leq \theta_{\text{crit}}^{(\text{RZ})}$) or a trivial extension of the COMP algorithm (for $\theta \geq \theta_{\text{crit}}^{(\text{RZ})}$). See Section III for details.
2. As $\rho \to 0$, we have $\theta_{\text{opt}}(\rho) \to 1/2$, and we recover the fact that the optimal performance is achieved by practical algorithms for $\theta \geq 1/2$ [12]. However, as soon as we increase the noise level even slightly, the achievability and converse match over a noticeably wider parameter range. For example, for $\rho = 0.001$, this is the case for $\theta \geq \theta_{\text{opt}}(\rho) = 0.3656$, and for $\rho = 0.1$ this widens to $\theta \geq \theta_{\text{opt}}(\rho) = 0.2119$.
3. As $\rho \to 0$, we have $\theta_{\text{crit}}^{(\text{RZ})} \to 1$, and so for any fixed θ the final case in (27) does not apply in this limit.

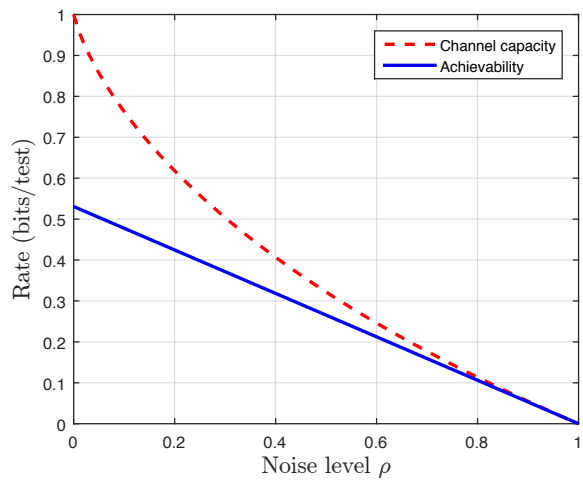

Figure 4: Reverse Z-channel model: Limiting achievable rate and algorithm-independent converse as $\theta \to 0$, plotted as a function of the noise level ρ.

Furthermore, as $\rho \to 0$, (18) gives $\frac{-\log \rho}{\kappa(\theta)} \to \frac{1-\theta}{\theta}$, and we recover the noiseless results (5)–(6).

4. Both the achievable rate and converse provide a curve which is continuous in θ. We can establish continuity at $\theta_{\text{crit}}^{(\text{RZ})}$ using the fact that $\kappa(\theta_{\text{crit}}^{(\text{RZ})}) = \frac{1}{\rho}$. This, in turn, follows because we can verify that $\rho^{\theta_{\text{crit}}^{(\text{RZ})}/(\theta_{\text{crit}}^{(\text{RZ})}-1)} e = \rho e^{1/\rho}$ (as (25) gives $\frac{\theta_{\text{crit}}^{(\text{RZ})}}{\theta_{\text{crit}}^{(\text{RZ})}-1} = \frac{1-\rho}{\rho \log \rho} + 1$ and in addition we have $\rho^{\frac{1-\rho}{\rho \log \rho}} = e^{1/\rho - 1}$), and the choice $\kappa = \frac{1}{\rho}$ makes $\kappa e^{-\kappa} = \left(\rho e^{1/\rho}\right)^{-1}$ in agreement with the definition $\kappa e^{-\kappa} = e^{-1} \rho^{\theta/(1-\theta)}$.

In the appendices, we provide two further claims pertaining to converse results under RZ noise:

1. (Appendix G) When the noisy DD algorithm is used in conjunction with Bernoulli testing, no rate higher than $\frac{1-\rho}{e \log 2}$ can be achieved. Therefore, one cannot hope to improve on the first case in (27) (nor on the other cases where an algorithm-independent converse holds) without moving to a different test design and/or a different decoding algorithm. To our knowledge, this result is new even when specialized to the noiseless case.
2. (Appendix H) In the limit $\theta \to 0$, the achievable rate approaches $\frac{1-\rho}{e \log 2}$, which is also the first-order term in the Z-channel capacity as $\rho \to 1$. Therefore, under the order of limits $n \to \infty$, $\theta \to 0$, and then $\rho \to 1$, the limiting behavior of the noisy DD rate cannot be improved on even by an adaptive algorithm (since the capacity-based converse holds even for adaptive algorithms [11]). We support this claim with the rate plot in Figure 4 for $\theta = 0$, where we observe nearly tight bounds for ρ close to one.

D. Z-Channel

Our rates for the Z-channel are written in terms of the following. For given values of ν and ρ, define $s = (1-\rho)e^{-\nu}/\rho > 0$ and for $\theta \neq \frac{1}{2}$ write the ratio

$$g(s,\theta) = \frac{1 + \frac{\theta}{2\theta-1} s}{(1+s)^{\theta/(2\theta-1)}}, \tag{29}$$

and introduce

$$\lambda := \lambda(\theta) = \begin{cases} W_0\left(-e^{-1} g(s,\theta)\right), & \text{for } \theta < \frac{1}{2}, \\ W_{-1}\left(-e^{-1} g(s,\theta)\right), & \text{for } \theta > \frac{1}{2}. \end{cases} \tag{30}$$

Moreover, define

$$\begin{aligned} \alpha^*(\theta) &= \rho \tfrac{s}{\log(1+s)} && \text{for } \theta = \tfrac{1}{2}, \\ \alpha^*(\theta) &= -\tfrac{\rho}{\lambda(\theta)}\left(1 + \tfrac{\theta s}{2\theta-1}\right) && \text{for } \theta \neq \tfrac{1}{2}. \end{aligned} \tag{31}$$

In the two cases $\theta < \frac{1}{2}$ and $\theta > \frac{1}{2}$, the value of $\frac{\theta}{2\theta-1}$ is below 0 and above 1, respectively. Since these are the two ranges of values of r for which the Bernoulli inequality $(1+rs) \leq (1+s)^r$ holds, in each case we can deduce that $g(s,\theta) \leq 1$, which implies that the Lambert functions in (30) are well-defined (see Figure 2).

Theorem 2. (Z-Channel) *For noisy group testing under Z-channel noise with parameter ρ, in the regime where $k \asymp p^\theta$ with $\theta \in (0,1)$, we have the following under Bernoulli testing with parameter $\nu > 0$:*

1. **[Achieveable rate]** *There exists a practical algorithm with error probability $P_{\text{e}} \to 0$ with*

$$\underline{R}^{(\text{Z})}(\theta,\rho) = \frac{(1-\theta)(1-\rho)\nu e^{-\nu}}{\theta \log 2} \times \min\left\{\frac{\theta}{2\theta-1}\left(\frac{\alpha^*(\theta)}{\alpha^*(1/2)} - 1\right), 1\right\} \quad \textit{if} \quad \theta \neq \frac{1}{2}, \tag{32a}$$

and

$$\underline{R}^{(\text{Z})}(\theta,\rho) = \frac{(1-\rho)\nu e^{-\nu}}{\log 2} \min\left\{\frac{1}{\log(1+s)} \log \frac{s}{\log(1+s)} - \frac{1}{\log(1+s)} + \frac{1}{s}, 1\right\} \quad \textit{if} \quad \theta = \frac{1}{2}. \tag{32b}$$

2. **[Converse]** *If $\rho < 1/2$, then no algorithm can achieve $P_{\text{e}} \to 0$ with a rate higher than*

$$\overline{R}^{(\text{Z})}(\theta,\rho) = \min\left\{C_Z(\rho), \frac{(1-\theta)(1-\rho)}{\theta e \log 2}\right\}. \tag{33}$$

Proof. See Section IV. □

These rates are illustrated for two different noise levels in Figure 5. The noisy DD algorithm used for the achievability part is described in Section IV. We proceed by giving some properties of these rates; see Remark 5 in the proof for more details.

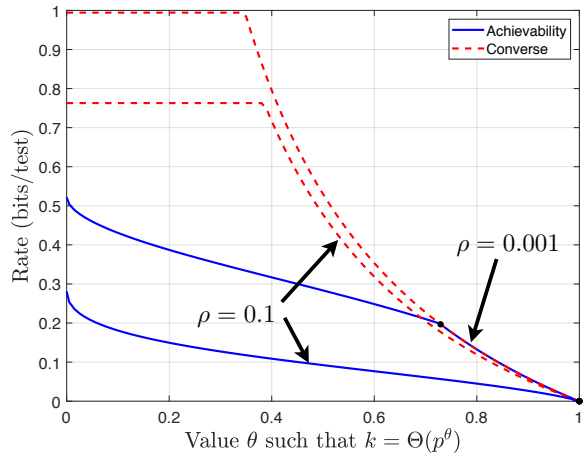

Figure 5: Achievable rate and algorithm-independent lower bound under Z-channel noise and Bernoulli testing. The dots indicate the threshold $\theta^{(\mathrm{Z})}_{\mathrm{crit}}$ (see Remark 3).

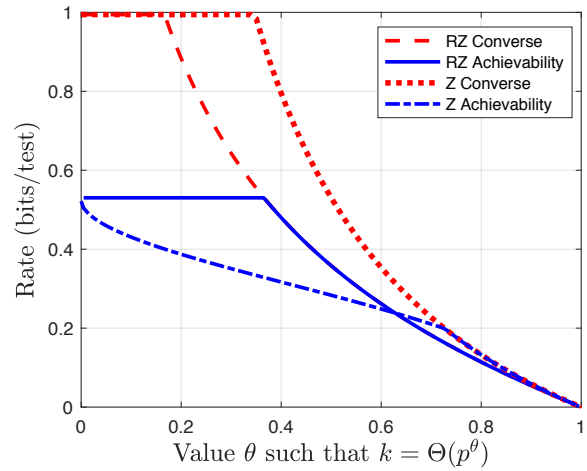

Figure 6: Comparison of rates under Z and reverse Z-channel noise.

Remark 3.

1. It can be shown that both the achievable and converse rates are continuous and non-increasing in θ.
2. There exists some $\theta^{(\mathrm{Z})}_{\mathrm{crit}}$ (which may be equal to 1) such that the first term achieves the $\min\{\cdot, 1\}$ in (32) if and only if $\theta \le \theta^{(\mathrm{Z})}_{\mathrm{crit}}$. For $\nu = 1$ and $\theta > \theta^{(\mathrm{Z})}_{\mathrm{crit}}$, the achievability and converse bounds coincide. As we see in Figure 5, this is indeed observed for small ρ; however, even for $\rho = 0.1$, the bounds do not match for any value of $\theta < 1$, due to the fact that $\theta^{(\mathrm{Z})}_{\mathrm{crit}} = 1$.
3. As $\theta \to 1$, we have $g(s, \theta) \to 1$, so that $\lambda(\theta) \to -1$ and $\alpha^*(\theta) \to \rho(1 + s)$. The implies that the first term inside the minimum in (32a) converges to $((1+s)\log(1+s) - s)/s$, which is less than 1 if and only if $s > 3.922$. In the case $\nu = 1$, this corresponds to the fact that the first term in (32) always gives the minimum (i.e., $\theta^{(\mathrm{Z})}_{\mathrm{crit}} = 1$) for $\rho < 0.0858$.
4. As $\rho \to 0$, the converse result (33) clearly tends to the noiseless converse result of (6). The corresponding argument for the achievability rate is more delicate. However, as described in Remark 5 in the proof, we indeed recover the noiseless achievable rate of (5) as $\rho \to 0$.
5. The achievability result includes the Bernoulli testing parameter $\nu > 0$, which can be optimized. Since s depends on ν, there appears to be no closed form expression for the optimal choice. However, our numerical findings suggest that the value $\nu = 1$ is near-optimal, particularly for small ρ. Comparing the achievability and converse parts, we see that $\nu = 1$ is certainly optimal when $\theta \ge \theta^{(\mathrm{Z})}_{\mathrm{crit}}$.

E. Comparison of the Channels

In Figure 6, we compare the rates for the Z and reverse Z-channel noise models. For θ close to one, we observe that the former is provably easier to handle: The Z achievability curve lies above the RZ converse curve. On the other hand, RZ noise appears to be easier to handle for small to moderate θ (though we cannot say this conclusively, as we have not yet verified whether the Z achievability curve is the best possible).

Some intuition behind this behavior can be obtained by noting that, like the noiseless version, the noisy DD algorithms first use negative tests to find a set of "possible defectives" (PD), and then estimate the defective set based on positive tests containing a single PD. The rate turns out to be dictated by the first step for small θ, and by the second step for large θ. Given that this is the case, the behavior in Figure 6 is to be expected:

- The first step is easier under RZ noise, since negative test outcomes are perfectly reliable.
- The second step is easier under Z noise, since positive test outcomes are perfectly reliable.

The fact that Z noise is preferable for θ close to one was also observed in the *adaptive* setting in [20].

F. General Binary Noise and Symmetric Noise

Our techniques can be extended to general binary channels (*cf.*, (7)), including the widely-considered symmetric model (*cf.*, (10)). However, for such channels, we have not yet proved a matching achievability and converse in any regime, other than the low-noise limit. We therefore defer our results on these models to Appendix I.

III. Proof of Theorem 1 (Reverse Z-Channel)

A. Achievability via Noisy DD and COMP

Recall the reverse Z-channel of Definition 2. We first describe a noisy DD algorithm for this model, exploiting the fact that if an item appears in a test with $Y = 0$ then

we can be certain that it is non-defective.

Noisy DD algorithm for RZ channel noise:

1. For each $j \in [p]$, let $N_{\mathrm{neg},j}$ be the number of negative tests in which item j is included. In the first step, we construct the following set of items that are definitely non-defective:
$$\widehat{\mathrm{ND}} = \left\{ j \in [p] \,:\, N_{\mathrm{neg},j} > 0 \right\}. \tag{34}$$
The remaining items, $\widehat{\mathrm{PD}} = [p] \setminus \widehat{\mathrm{ND}}$, are called "possibly defective".
2. For each $j \in \widehat{\mathrm{PD}}$, let $N'_{\mathrm{pos},j}$ be the number of positive tests that include item j and no other item from $\widehat{\mathrm{PD}}$. In the second step, we fix a constant $\beta \in (\rho, 1)$, and estimate the defective set as follows:
$$\widehat{S} = \left\{ j \in \widehat{\mathrm{PD}} \,:\, N'_{\mathrm{pos},j} \geq \frac{\beta n \nu e^{-\nu}}{k} \right\}. \tag{35}$$

In addition to this noisy variation of DD, we can directly apply the noiseless COMP algorithm (e.g., see [23]) to the reverse Z-channel model. As in the noiseless case, we know that even if reverse Z-channel noise is present, an item appearing in a negative test is definitive proof that it is not defective. Therefore, we can consider taking $\widehat{\mathrm{PD}}$ above to be the estimate of S. We analyze the performance of this algorithm using essentially the same argument as in [23]. We wish to ensure that each non-defective item has some test where it is tested with no defective present, and that the resulting test outcome is not changed by the reverse Z-channel (we say that the non-defective item *survives* this test).

Lemma 1. (COMP under RZ noise) *Consider the reverse Z-channel noisy group testing setup with parameter $\rho \in (0,1)$, number of defectives $k \asymp p^\theta$ (where $\theta \in (0,1)$), and i.i.d. Bernoulli testing with parameter $\nu > 0$. Under the COMP algorithm, we have $P_{\mathrm{e}} \to 0$ as long as $n \geq n_{\mathrm{COMP}}(1+\eta)$ for arbitrarily small $\eta > 0$, where*
$$n_{\mathrm{COMP}} = \frac{1}{(1-\rho)\nu e^{-\nu}} k \log p. \tag{36}$$

Proof. The probability that any particular non-defective item (is included in and) survives any particular test is $(1-\rho)\nu/k(1-\nu/k)^k$. Using $n = \gamma k \log p$ tests, we can control P_{e} using the union bound by
$$P_{\mathrm{e}} \leq \mathbb{P}\left(\bigcup_{j \in S^c} \{\text{item } j \text{ doesn't survive any test}\} \right) \tag{37}$$
$$\leq p\mathbb{P}\left(\text{item } j \text{ doesn't survive any test}\right) \tag{38}$$
$$= p\left(1 - (1-\rho)\frac{\nu}{k}\left(1-\frac{\nu}{k}\right)^k\right)^n \tag{39}$$
$$\leq p \exp\left(\left(-(1-\rho)\nu(1-\nu/k)^k\gamma\right)\log p\right) \tag{40}$$
$$\leq \exp\left(\left(1-(1-\rho)\nu(1-\nu/k)^k\gamma\right)\log p\right), \tag{41}$$
where (40) uses $1 - z \leq e^{-z}$. Since $\left(1-\frac{\nu}{k}\right)^k = e^{-\nu}(1+o(1))$, we deduce that if $n \geq n_{\mathrm{COMP}}(1+\eta)$ (or equivalently $(1-\rho)\nu e^{-\nu}\gamma \geq 1+\eta$), then $(1-\rho)\nu(1-\nu/k)^k\gamma \geq 1+\eta/2$ for k sufficiently large, so $P_{\mathrm{e}} \leq p^{-\eta/2}$, which converges to 0. □

The main step towards proving the achievability part of Theorem 1 is to establish the following.

Theorem 3. (NDD under RZ noise) *Consider the reverse Z-channel noisy group testing setup with parameter $\rho \in (0,1)$, number of defectives $k \asymp p^\theta$ (where $\theta \in (0,1)$), and i.i.d. Bernoulli testing with parameter $\nu > 0$. For any $\beta \in (\rho, 1)$ and $\xi \in (0, \theta)$, the noisy DD algorithm achieves $P_{\mathrm{e}} \to 0$ as long as $n \geq n_{\mathrm{DD}}(1+\eta)$ for arbitrarily small $\eta > 0$, where*
$$n_{\mathrm{DD}} = \max\left\{ n_1^{(\mathrm{ND})}, n_2^{(\mathrm{D})}, n_2^{(\mathrm{ND})} \right\}, \tag{42}$$
and where
$$n_1^{(\mathrm{ND})} = \frac{1-\xi}{(1-\rho)\nu e^{-\nu}} \cdot k \log p, \tag{43}$$
$$n_2^{(\mathrm{D})} = \frac{1}{\nu e^{-\nu} D_1(\beta)} \cdot k \log k, \tag{44}$$
$$n_2^{(\mathrm{ND})} = \frac{\xi}{\nu e^{-\nu} \rho D_1(\beta/\rho)} \cdot k \log p. \tag{45}$$

Once this result and Lemma 1 are in place, proving the achievability of the rate (27) boils down to algebraic manipulations. Since these are somewhat tedious, they are deferred to Appendix B. In the remainder of this subsection, we focus on the proof of Theorem 3.

Let N_{neg} and N_{pos} respectively denote the number of negative and positive tests. Both of these quantities follow a binomial distribution; the probability of a given test i being positive is given by
$$\mathbb{P}[Y_i = 1] = \left(1 - \left(1-\frac{\nu}{k}\right)^k\right) + \rho\left(1-\frac{\nu}{k}\right)^k \tag{46}$$
$$= \left((1-e^{-\nu}) + \rho e^{-\nu}\right)(1+o(1)), \tag{47}$$
since $\left(1-\frac{\nu}{k}\right)^k \to e^{-\nu}$ as $k \to \infty$, and we similarly have
$$\mathbb{P}[Y_i = 0] = (1-\rho)e^{-\nu} \cdot (1+o(1)). \tag{48}$$
Hence, and using (19)–(20), we have with probability approaching one that
$$N_{\mathrm{neg}} = n \cdot (1-\rho)e^{-\nu} \cdot (1+o(1)), \tag{49}$$
$$N_{\mathrm{pos}} = n \cdot \left((1-e^{-\nu}) + \rho e^{-\nu}\right) \cdot (1+o(1)). \tag{50}$$
It will be useful to split the N_{pos} positive tests into two types: those that contain a defective item and whose noise does not flip the outcome, and those that contain no defective items but whose noise flips the outcome. The

number of such tests are denoted by $N_{\rm pos}^{\rm (D)}$ and $N_{\rm pos}^{\rm (ND)}$, respectively. A given test falls into the first category with probability $(1-e^{-\nu})(1+o(1))$, and the second category with probability $\rho e^{-\nu}(1+o(1))$. Therefore, the concentration bounds (19)–(20) yield

$$N_{\rm pos}^{\rm (D)} = n\cdot(1-e^{-\nu})\cdot(1+o(1)), \tag{51}$$

$$N_{\rm pos}^{\rm (ND)} = n\cdot \rho e^{-\nu}\cdot(1+o(1)). \tag{52}$$

with probability approaching one.

Throughout the remainder of the analysis, we implicitly condition on the defective set S taking a fixed value, say $S=\{1,\dots,k\}$. By the symmetry of the random test design, the conditional error probability is the same for any such realization of cardinality k.

Analysis of First Step. Since negative tests can never include a defective item for the reverse Z-channel, the set $\widehat{\rm PD}=[p]\setminus\widehat{\rm ND}$ contains all of the defective items. We proceed by establishing a sufficient condition such that with high probability, it also contains at most p^{ξ} non-defectives, for some constant $\xi\in(0,\theta)$. We denote the complement of this event by $P_{\rm e,1}^{\rm (ND)}$, where the subscript and superscript respectively denote the step number and the consideration of non-defective items.

Analysis of non-defective items. Let $P_{\rm e,1}^{\rm (ND)}(n_{\rm neg})$ be defined similarly to $P_{\rm e,1}^{\rm (ND)}$, but conditioned on $N_{\rm neg}$ taking a given value $n_{\rm neg}$. It suffices to establish that $P_{\rm e,1}^{\rm (ND)}(n_{\rm neg})\to 0$ for all $n_{\rm neg}$ satisfying the concentration bound (49), since this bound holds with probability approaching one.

Since the test outcomes depend only on the defective items, one can envision the non-defective items as being placed in each test with probability $\frac{\nu}{k}$ *after* the test outcomes have been produced. As a result, given $N_{\rm neg}=n_{\rm neg}$, the number of negative tests $N_{{\rm neg},j}$ including a given item $j\notin S$ is distributed as $(N_{{\rm neg},j}|n_{\rm neg})\sim {\rm Binomial}\big(n_{\rm neg},\frac{\nu}{k}\big)$. Hence, for any $j\notin S$, we have

$$\mathbb{P}[j\notin\widehat{\rm ND}\,|\,n_{\rm neg}] = \mathbb{P}\big[N_{{\rm neg},j}=0\,\big|\,n_{\rm neg}\big] \tag{53}$$

$$= \left(1-\frac{\nu}{k}\right)^{n_{\rm neg}} \tag{54}$$

$$\le e^{-\frac{n_{\rm neg}\nu}{k}}, \tag{55}$$

where we have applied $1-\zeta\le e^{-\zeta}$. As a result, letting G denote the number of non-defective items in $\widehat{\rm PD}=[p]\setminus\widehat{\rm ND}$, we find that $\mathbb{E}[G\,|\,n_{\rm neg}]$ is upper bounded by $p-k$ times the right-hand side of (55). Applying Markov's inequality, we obtain

$$\mathbb{P}[G\ge p^{\xi}\,|\,n_{\rm neg}]\le p^{1-\xi}e^{-\frac{n_{\rm neg}\nu}{k}}. \tag{56}$$

Since $n_{\rm neg}$ satisfies (49), we find that we can achieve $P_{\rm e,1}^{\rm (ND)}\to 0$ under the condition

$$n\ge\left(\frac{1-\xi}{(1-\rho)\nu e^{-\nu}}\cdot k\log p\right)(1+o(1)). \tag{57}$$

Analysis of Second Step. Recall that the final estimate includes all $j\in\widehat{\rm PD}$ for which $N'_{{\rm pos},j}\ge\frac{\beta n\nu e^{-\nu}}{k}$, where $N'_{{\rm pos},j}$ is the number of tests containing item j and no other item from $\widehat{\rm PD}$. To characterize the distribution of $N'_{{\rm pos},j}$, we make use of a multinomial conditioning argument analogous to that used in the noiseless DD algorithm [12]. Since this is used multiple times throughout the paper, we first state the relevant result in generic notation.

Lemma 2. [12, Lemma C.1] *Fix a positive integer m, and let (W_0,W_1,W_2) have a multinomial distribution with m trials and probabilities (r_0,r_1,r_2). Associate an observation $(W_0,W_1,W_2)=(w_0,w_1,w_2)$ with an unordered list of m class labels (class 0, 1, or 2), and suppose that each label in class 1 is independently changed to some sub-class $1'$ with probability $\gamma\in[0,1]$, and to some sub-class $1''$ with probability $1-\gamma$ (where γ may depend on w_0). Then, conditioned on $W_0=w_0$, the corresponding random variables (W_1',W_1'',W_2) counting the transformed class labels have a multinomial distribution with $m-w_0$ trials and the following probability parameters:*

$$\left(\frac{r_1\gamma}{1-r_0},\frac{r_1(1-\gamma)}{1-r_0},\frac{r_2}{1-r_0}\right). \tag{58}$$

Analysis of defective items. To apply Lemma 2, we first fix a defective $j\in S$ and consider the triplet $(N_{\rm neg},\widetilde{N}_{{\rm pos},j},N_{\rm other})$, where:

- $N_{\rm neg}$ is the number of negative tests; recall from (48) that the probability of a negative test is $q_{\rm neg}=\big((1-\rho)e^{-\nu}\big)(1+o(1))$.
- $\widetilde{N}_{{\rm pos},j}$ is the number of tests containing j but no other defective item; the probability of a given test satisfying this condition is $\widetilde{q}_j=\frac{\nu}{k}\big(1-\frac{\nu}{k}\big)^{k-1}=\big(\frac{\nu e^{-\nu}}{k}\big)(1+o(1))$.
- $N_{\rm other}$ is the number of remaining tests, with associated probability $q_{\rm other}=1-q_{\rm neg}-\widetilde{q}_j$.

Hence, $(N_{\rm neg},\widetilde{N}_{{\rm pos},j},N_{\rm other})$ has a multinomial distribution with n trials and parameters $(q_{\rm neg},\widetilde{q}_j,q_{\rm other})$.

We now consider conditioning on the negative tests, and consequently on $N_{\rm neg}=n_{\rm neg}$, $\widehat{\rm PD}=\widehat{\rm pd}$, and $G=g$, which are determined from the first step using only these tests. We assume that $\widehat{\rm pd}$ contains all defective items and $g\le p^{\xi}$ non-defectives, and that $n_{\rm neg}$ satisfies the concentration bound (49); recall that these events all occur with probability approaching one.

Conditioned on $(\widehat{\rm pd},g,n_{\rm neg})$, we consider "splitting" the $\widetilde{N}_{{\rm pos},j}$ tests mentioned above into two sub-classes with counts $(N'_{{\rm pos},j},\widetilde{N}_{{\rm pos},j}-N'_{{\rm pos},j})$. By the definitions of $\widetilde{N}_{{\rm pos},j}$ and $N'_{{\rm pos},j}$, each test will fall in the second sub-class if it contains a non-defective from $\widehat{\rm pd}$, and in

the first class otherwise. Hence, the associated conditional probability of falling in the first sub-class is[3]

$$\gamma = \left(1 - \frac{\nu}{k}\right)^g = 1 - o(1), \tag{59}$$

where we used the fact that $g = o(k)$ (since $g \le p^\xi$ with $\xi < \theta$, whereas $k \asymp p^\theta$).

Combining the above observations and applying Lemma 2, we find that the joint distribution of $(N'_{\mathrm{pos},j}, \widetilde{N}_{\mathrm{pos},j} - N'_{\mathrm{pos},j}, N_{\mathrm{other}})$ given $(\widehat{\mathrm{pd}}, g, n_{\mathrm{neg}})$ is multinomial with $n - n_{\mathrm{neg}}$ trials, with the first probability parameter being

$$q'_{\mathrm{pos},j} = \frac{\widetilde{q}_j \gamma}{1 - q_{\mathrm{neg}}} \tag{60}$$

$$= \left(\frac{\nu e^{-\nu}}{k(1 - q_{\mathrm{neg}})}\right)(1 + o(1)) \tag{61}$$

by the above calculations of $\widetilde{q}_j$ and γ. Since the marginal of a multinomial distribution is binomial and $n - n_{\mathrm{neg}} = n(1 - q_{\mathrm{neg}})(1 + o(1))$ (see (49)), we deduce that

$$(N'_{\mathrm{pos},j} \,|\, \widehat{\mathrm{pd}}, g, n_{\mathrm{neg}}) \sim \mathrm{Binomial}\bigg(n(1-q_{\mathrm{neg}})(1+o(1)), \frac{\nu e^{-\nu}}{k(1 - q_{\mathrm{neg}})}(1 + o(1))\bigg). \tag{62}$$

Observe that the mean $\tilde{\mu}_1^{(\mathrm{D})}$ of this distribution satisfies

$$\tilde{\mu}_1^{(\mathrm{D})} = \left(\frac{n}{k}\nu e^{-\nu}\right)(1 + o(1)). \tag{63}$$

Recall from (35) that a given item j is included in the final estimate $\widehat{S}$ if $N'_{\mathrm{pos},j} \ge \frac{\beta n \nu e^{-\nu}}{k}$. Under the definition $\epsilon_2^{(\mathrm{D})} = 1 - \beta$, we find that the threshold $\frac{\beta n \nu e^{-\nu}}{k}$ equals $\frac{n}{k}\nu e^{-\nu}(1 - \epsilon_2^{(\mathrm{D})})$, and hence, the probability that a given defective item j is incorrectly excluded from $\widehat{S}$ satisfies

$$\mathbb{P}[j \notin \widehat{S} \,|\, \widehat{\mathrm{pd}}, g, n_{\mathrm{neg}}]$$

$$= \mathbb{P}\left[N'_{\mathrm{pos},j} < \frac{n}{k}\nu e^{-\nu}(1 - \epsilon_2^{(\mathrm{D})}) \,\middle|\, \widehat{\mathrm{pd}}, g, n_{\mathrm{neg}}\right] \tag{64}$$

$$\le \exp\left(-\frac{n}{k}\nu e^{-\nu} \cdot D_1(1 - \epsilon_2^{(\mathrm{D})}) \cdot (1 + o(1))\right) \tag{65}$$

by (62)–(63) and the concentration bound in (19).

By the union bound, the probability of there existing some $j \in S$ failing to be included in $\widehat{S}$ is at most k times the right-hand side of (65). By re-arranging, we deduce that $P_{\mathrm{e},2}^{(\mathrm{D})} \to 0$ as $p \to \infty$ under the condition

$$n \ge \left(\frac{1}{\nu e^{-\nu}} \cdot \frac{1}{D_1(\beta)} \cdot k \log k\right)(1 + o(1)), \tag{66}$$

[3]In Lemma 2 we only condition on $W_0 = w_0$, which plays the role of $N_{\mathrm{neg}} = n_{\mathrm{neg}}$ here. Since the tests are independent and the conditioning on $\widehat{\mathrm{PD}} = \widehat{\mathrm{pd}}$ and $G = g$ only depends on the negative tests, this additional conditioning does not affect the conditional joint distribution of $(\widetilde{N}_{\mathrm{pos},j}, N_{\mathrm{other}})$ given $N_{\mathrm{neg}} = n_{\mathrm{neg}}$; it only affects the probability γ appearing in Lemma 2.

where we have substituted the choice $\epsilon_2^{(\mathrm{D})} = 1 - \beta$.

Analysis of non-defective items. To characterize the probability of a non-defective incorrectly being included in the final estimate, we apply a similar argument to the one following Lemma 2. Due to the level of similarity, we omit some details and focus on the main differences.

We consider the triplet $(N_{\mathrm{neg}}, \widetilde{N}_{\mathrm{pos},j}, N_{\mathrm{other}})$ defined in the same way as above, but now with $\widetilde{N}_{\mathrm{pos},j}$ being the number of *positive* tests containing a given *non-defective* j and none of the defective items. The associated probability is $\widetilde{q}_j = \rho\frac{\nu}{k}\big(1 - \frac{\nu}{k}\big)^k = \big(\frac{\rho\nu e^{-\nu}}{k}\big)(1 + o(1))$, and the probability $q_{\mathrm{other}} = 1 - q_{\mathrm{neg}} - \widetilde{q}_j$ changes accordingly.

We again condition on $(\widehat{\mathrm{pd}}, g, n_{\mathrm{neg}})$, and consider the splitting the $\widetilde{N}_{\mathrm{pos},j}$ tests into two sub-classes with counts $(N'_{\mathrm{pos},j}, \widetilde{N}_{\mathrm{pos},j} - N'_{\mathrm{pos},j})$. The same argument as (59) gives $\gamma = 1 - o(1)$, and we obtain the following analog of (62):

$$(N'_{\mathrm{pos},j} \,|\, n_{\mathrm{neg}}, \widehat{\mathrm{pd}}, g) \sim \mathrm{Binomial}\bigg(n(1-q_{\mathrm{neg}})(1+o(1)), \frac{\rho\nu e^{-\nu}}{k(1 - q_{\mathrm{neg}})}(1 + o(1))\bigg), \tag{67}$$

with the mean satisfying

$$\tilde{\mu}_1^{(\mathrm{ND})} = \left(\frac{n}{k}\rho\nu e^{-\nu}\right)(1 + o(1)). \tag{68}$$

Defining $\epsilon_2^{(\mathrm{ND})} = \frac{\beta - \rho}{\rho}$, we find that the threshold $\frac{\beta n \nu e^{-\nu}}{k}$ equals $\frac{n}{k}\rho\nu e^{-\nu}(1 + \epsilon_2^{(\mathrm{ND})})$, and hence, the probability of $j \notin S$ incorrectly being included in $\widehat{S}$ satisfies

$$\mathbb{P}[j \in \widehat{S} \,|\, \mathcal{E}_1^c, g, n_{\mathrm{pos}}^{(\mathrm{D})}]$$

$$= \mathbb{P}\left[N'_{\mathrm{pos},j} \ge \frac{n}{k}\rho\nu e^{-\nu}(1 + \epsilon_2^{(\mathrm{ND})}) \,\middle|\, \mathcal{E}_1^c, g, n_{\mathrm{pos}}^{(\mathrm{D})}\right] \tag{69}$$

$$\le \exp\left(-\frac{n}{k}\rho\nu e^{-\nu} \cdot D_1(1 + \epsilon_2^{(\mathrm{ND})}) \cdot (1 + o(1))\right) \tag{70}$$

by (67)–(68) and the concentration bound in (20).

By the union bound, the probability of any $j \in \widehat{\mathrm{PD}} \setminus S$ incorrectly being included in $\widehat{S}$ is at most $g \le p^\xi$ times the right-hand side of (70). By re-arranging, we deduce that $P_{\mathrm{e},2}^{(\mathrm{ND})} \to 0$ as $p \to \infty$ under the condition

$$n \ge \left(\frac{\xi}{\rho\nu e^{-\nu}} \cdot \frac{1}{D_1(\beta/\rho)} \cdot k \log p\right)(1 + o(1)), \tag{71}$$

where we have substituted $\epsilon_2^{(\mathrm{ND})} = \frac{\beta - \rho}{\rho} = \frac{\beta}{\rho} - 1$.

Wrapping up. Combining the conditions in (57), (66), and (71), we deduce Theorem 3.

B. Algorithm-Independent Converse

We will prove a converse that holds for all algorithms under Bernoulli testing. Since S is assumed to be uniform

over the subsets of cardinality k, the decoding rule that minimizes $P_{\rm e}$ is maximum likelihood:

$$\widehat{S}_{\rm ML}(\mathbf{y},\mathbf{X}) = \underset{S\,:\,|S|=k}{\arg\max}\ \mathbb{P}[\mathbf{Y}=\mathbf{y}|\mathbf{X},S]. \tag{72}$$

Hence, it suffices to lower bound the number of tests required for vanishing error probability under this decoding rule. We first define two families of random variables:

1. Given a defective item $i \in S$, as in [12], we write M_i for the number of tests containing defective item i but no other defective.
2. We consider the set $\mathcal{T}$ of intruding possible defectives, defined as follows: $j \in \mathcal{T}$ if item j is non-defective and does not appear in any negative tests. Given an intruding possible defective $j \in \mathcal{T}$, we write N_j for the number of false positive tests containing j.

Let $N_{\rm pos}^{\rm (ND)}$ denote the overall total number of false positive tests introduced by the reverse Z-channel. The key idea is to look for a defective item $i \in S$ with $M_i \le c$ (for some c to be determined), and an intruding possible defective item $j \in \mathcal{T}$ with $N_j > c$. If we can find such items, then the set $S \setminus \{i\} \cup \{j\}$ explains the outcomes better than S (since it requires at most $c + (N_{\rm pos}^{\rm (ND)} - N_j) < N_{\rm pos}^{\rm (ND)}$ errors to be made by the reverse Z-channel, yielding a higher likelihood assuming $\rho < 1/2$). As a result, the optimal maximum likelihood (ML) algorithm will make a mistake.

Suppose that $n = \gamma k \log p$, which equates to a rate of $\frac{1-\theta}{\gamma \log 2}$. We first argue that there exists a defective item that is not the unique one in too many tests.

Lemma 3. *Fix $\gamma > 0$ and $\Phi < 1$, and let $c = \Phi\nu e^{-\nu}\gamma\log p$ and $n = \gamma k \log p$. If $\theta > D_1(\Phi)\nu e^{-\nu}\gamma$, then with probability approaching one there exists a defective item $i \in S$ such that $M_i \le c$.*

Proof. As in [12], the M_i (together with other random variables corresponding to there being no defectives and multiple defectives in a test) are jointly multinomially distributed. Writing $r = \nu(1-\nu/k)^{k-1} \le \nu e^{-\nu}$, the marginal distribution of each M_i is $\mathrm{Binomial}(n, r/k)$, so $\mathbb{E}[M_i] = r\gamma\log p$. In addition, we have

$$\begin{aligned}\mathbb{P}\left[M_i \le c\right] &= \mathbb{P}\left[M_i \ge \Phi\nu e^{-\nu}\gamma\log p\right]\\ &\le \mathbb{P}\left[M_i \ge \Phi r\gamma\log p\right].\end{aligned} \tag{73}$$

Using the concentration bound (19) (with $N = n = \gamma k\log p$, $q = r/k$, and $\epsilon = 1-\Phi$, also implying $Nq = r\gamma\log p$), we deduce that

$$\mathbb{P}\left[M_i \le c\right] \le p^{-D_1(\Phi)r\gamma}. \tag{74}$$

Next, we use the key fact (see [45, Section 3.1]) that multinomial random variables satisfy the so-called negative association property, which implies for any c that [45, Property P3]

$$\begin{aligned}\mathbb{P}\left[\min_{i\in S} M_i \ge c+1\right] &= \mathbb{P}\left[\bigcap_{i\in S}\{M_i \ge c+1\}\right]\\ &\le \prod_{i\in S}\mathbb{P}[M_i \ge c+1]\\ &= \left(1-\mathbb{P}[M_i \le c]\right)^k.\end{aligned} \tag{75}$$

Using (73) and (75), we deduce that

$$\begin{aligned}\mathbb{P}\left[\min_{i\in S} M_i \le c\right] &= 1-\mathbb{P}\left[\min_{i\in S} M_i \ge c+1\right] &(76)\\ &\ge 1-\left(1-\mathbb{P}[M_i \le c]\right)^k &(77)\\ &\ge 1-\exp(-k\mathbb{P}[M_i \le c]) &(78)\\ &= 1-\exp(-p^{\theta - D_1(\Phi)r\gamma}). &(79)\end{aligned}$$

Recalling that $r = \nu(1-\nu/k)^{k-1} = \big(\nu e^{-\nu}\big)(1+o(1))$, we deduce that if $\theta > D_1(\Phi)\nu e^{-\nu}\gamma$, then $\mathbb{P}\left[\min_{i\in S} M_i \le c\right] \to 1$. □

Similarly, we can find an intruding possible defective that appears in sufficiently many false positive tests.

Lemma 4. *Fix $\gamma > 0$ and $\Phi > \rho$, and let $c = \Phi\nu e^{-\nu}\gamma\log p$ and $n = \gamma k\log p$. If $\gamma\nu e^{-\nu}(D_\rho(\Phi)+1-\rho) < 1$, then with probability tending to 1 there exists a non-defective item j such that $N_j > c$.*

Proof. The probability of a particular test containing no defectives is $(1-\nu/k)^k$; hence, the probability that a test is negative is $(1-\rho)(1-\nu/k)^k$ and the probability that a test is a false positive is $\rho(1-\nu/k)^k$. The total number of such tests (which we refer to as $N_{\rm neg}$ and $N_{\rm pos}^{\rm (ND)}$ respectively) both have a binomial marginal distribution, due to the independence among tests. Hence, since $(1-\nu/k)^k \to e^{-\nu}$ from below, using concentration results of the form (19) and (20), for arbitrarily small $\epsilon' > 0$ we may assume that $N_{\rm neg} \le \mathbb{E}[N_{\rm neg}](1+\epsilon') \le n(1-\rho)e^{-\nu}(1+\epsilon')$ and $N_{\rm pos}^{\rm (ND)} \ge \mathbb{E}[N_{\rm pos}^{\rm (ND)}](1-\epsilon'/2) \ge n\rho e^{-\nu}(1-\epsilon')$ for p sufficiently large.

Conditioned on any such $N_{\rm neg} = n_{\rm neg}$, the probability of a given $j \in S^c$ belonging to $\mathcal{T}$ (i.e., being an intruding non-defective) is

$$\begin{aligned}&(1-\nu/k)^{n_{\rm neg}}\\ &\ge (1-\nu/k)^{ne^{-\nu}(1-\rho)(1+\epsilon')} &(80)\\ &= \exp\left(-\nu e^{-\nu}(1-\rho)(1+\epsilon')\gamma\log p(1-o(1))\right), &(81)\end{aligned}$$

where we recall that $n = \gamma k\log p$. Hence, again applying binomial concentration (with $p-k$ trials), we have with probability tending to one that the expected number of intruding possible defectives satisfies $|\mathcal{T}| \ge p^\tau/2$, where

$$\tau = 1-\nu e^{-\nu}(1-\rho)(1+2\epsilon')\gamma. \tag{82}$$

We seek to find a possible defective item lying in at least c of the $N_{\rm pos}^{(\rm ND)}$ false positive tests. Since (conditioned on $N_{\rm pos}^{(\rm ND)} = n_{\rm pos}^{(\rm ND)}$) it holds that $N_j \sim \mathrm{Binomial}(n_{\rm pos}^{(\rm ND)}, \nu/k)$, and since by assumption $n_{\rm pos}^{(\rm ND)} \geq n\rho e^{-\nu}(1-\epsilon')$, we know that N_j is stochastically dominated by $N_j^\dagger$, defined to be $\mathrm{Binomial}(n\rho e^{-\nu}(1-\epsilon'), \nu/k)$. Since $\rho < \Phi$, by a similar argument to the proof of Lemma 3, we have

$$\mathbb{P}\left[\max_{j\in\mathcal{T}} N_j > c\right] \geq \mathbb{P}\left[\max_{j\in\mathcal{T}} N_j^\dagger > c\right] \geq 1 - \exp\left(-\frac{1}{2}p^\tau \mathbb{P}[N_j^\dagger > c]\right). \quad (83)$$

Taking $N = ne^{-\nu}\rho$ and $q = \nu/k$ and $\epsilon = \Phi/\rho - 1$, we know from the discussion following (23) that (20) has a matching lower bound up to a $p^{-\epsilon'}$ pre-factor. Combining this observation with (15) gives

$$\begin{aligned} &p^\tau \mathbb{P}[N_j^\dagger > c] \\ &\geq p^{\tau-\epsilon'} \exp\left(-\gamma\nu e^{-\nu}\rho \log p D_1(\Phi/\rho)\right) && (84)\\ &= \exp\left(\log p\left(\tau - \epsilon' - \gamma e^{-\nu}\nu D_\rho(\Phi)\right)\right) && (85)\\ &= \exp\Big(\log p\big(1 - \epsilon' - \gamma\nu e^{-\nu}(D_\rho(\Phi) \\ &\qquad + (1-\rho)(1+2\epsilon')\big)\Big), && (86)\end{aligned}$$

where (85) uses (15), and (86) substitutes the definition of τ in (82). Hence, if $\gamma\nu e^{-\nu}(D_\rho(\Phi) + 1 - \rho) < 1$, then we can choose ϵ' sufficiently small such that in (83), the probability $\mathbb{P}\left[\max_{j\in\mathcal{T}} N_j > c\right]$ tends to one. □

We now put Lemmas 3 and 4 together. Suppose that there exists $\gamma > 0$ and $\Phi \in (\rho, 1)$ such that the rate $R = \frac{1-\theta}{\gamma \log 2}$ satisfies

$$\frac{1-\theta}{\gamma\log 2} \geq \max\left\{\left(\frac{(1-\theta)\nu e^{-\nu}}{\log 2}\right)\frac{D_1(\Phi)}{\theta}, \left(\frac{(1-\theta)\nu e^{-\nu}}{\log 2}\right)(D_\rho(\Phi) + 1 - \rho)\right\}. \quad (87)$$

Then, with probability approaching one, by Lemma 3 there exists a defective item i in fewer than c tests with no other defective, and by Lemma 4 a non-defective item j which appears in more than c false positive tests. In this case, the set $S \setminus \{i\} \cup \{j\}$ is preferred by the ML decoder to the true defective set S, and an error occurs.

If we perform too few tests, then γ will be small, and so $\frac{1}{\gamma}$ will be large, meaning (87) eventually holds. We would like to find the largest value for γ that allows such a Φ to exist. Similarly to the achievability proof, this final step amounts to somewhat tedious algebra, so the details are deferred to Appendix C.

IV. Proof of Theorem 2 (Z-Channel)

While the achievability and converse proofs below use similar ideas to those of the previous section, the details differ enough that we consider it worthwhile to include most steps.

A. Achievability via Noisy DD

Recall the Z-channel of Definition 2. We first describe a noisy DD algorithm for this model.

Noisy DD algorithm for Z-channel noise:

1. For each $j \in [p]$, let $N_{{\rm neg},j}$ be the number of negative tests in which item j is included. In the first step, we fix a constant $\alpha \in (\rho, 1)$ and construct the following set of items that are believed to be non-defective:
$$\widehat{\rm ND} = \left\{j \in [p] \,:\, N_{{\rm neg},j} \geq \frac{\alpha n \nu}{k}\right\}. \quad (88)$$
The remaining items, $\widehat{\rm PD} = [p] \setminus \widehat{\rm ND}$, are believed to be "possibly defective".
2. For each $j \in \widehat{\rm PD}$, let $N'_{{\rm pos},j}$ be the number of positive tests that include item j and no other item from $\widehat{\rm PD}$. In the second step, we estimate the defective set as follows:
$$\widehat{S} = \left\{j \in \widehat{\rm PD} \,:\, N'_{{\rm pos},j} > 0\right\}. \quad (89)$$

Since positive tests must contain a defective item under the Z-channel model, we deduce that as long as the first step is correct (in the sense that $S \subseteq \widehat{\rm PD}$), the second step will never add a defective item to $\widehat{S}$.

Theorem 4. *Consider the Z-channel noisy group testing setup with parameter $\rho \in (0,1)$, number of defectives $k \asymp p^\theta$ (where $\theta \in (0,1)$), and i.i.d. Bernoulli testing with parameter $\nu > 0$. For any $\alpha \in (\rho, 1)$, the noisy DD algorithm achieves $P_{\rm e} \to 0$ as long as*

$$n \geq \max\left\{n_1^{(\rm D)}, n_1^{(\rm ND)}, n_2^{(\rm D)}\right\}(1+\eta) \quad (90)$$

for arbitrarily small $\eta > 0$, where defining $\zeta = e^{-\nu} + \rho(1-e^{-\nu})$, we have

$$n_1^{(\rm D)} = \frac{1}{\nu\rho D_1(\alpha/\rho)} \cdot k\log k, \quad (91)$$

$$n_1^{(\rm ND)} = \frac{1}{\nu\zeta D_1(\alpha/\zeta)} \cdot k\log\frac{p}{k}, \quad (92)$$

$$n_2^{(\rm D)} = \frac{1}{(1-\rho)\nu e^{-\nu}} \cdot k\log k. \quad (93)$$

Similarly to the previous section, the remaining details in proving the achievability part of Theorem 2 using Theorem 4 amount to tedious algebra. We therefore defer these to Appendix D, and focus on proving Theorem 4.

Let $N_{\rm neg}$ and $N_{\rm pos}$ respectively denote the number of negative and positive tests. Both of these quantities follow a binomial distribution; the probability of a given test i being positive is given by

$$\mathbb{P}[Y_i = 1] = (1-\rho)\left(1 - \left(1 - \frac{\nu}{k}\right)^k\right) \quad (94)$$

$$= \Big((1-\rho)(1-e^{-\nu})\Big)(1+o(1)), \quad (95)$$

since $\big(1-\frac{\nu}{k}\big)^k \to e^{-\nu}$ as $k\to\infty$, and we similarly have

$$\mathbb{P}[Y_i=0] = \Big(e^{-\nu}+\rho(1-e^{-\nu})\Big)(1+o(1)). \tag{96}$$

Hence, and using (19)–(20), we have with probability approaching one that

$$N_{\rm neg} = n\Big(e^{-\nu}+\rho(1-e^{-\nu})\Big)(1+o(1)), \tag{97}$$

$$N_{\rm pos} = n\Big((1-\rho)(1-e^{-\nu})\Big)(1+o(1)). \tag{98}$$

We again henceforth condition on the defective set S taking a fixed value, say $S=\{1,\dots,k\}$.

Analysis of First Step. For the first step, we seek to ensure that, with probability approaching one, $\widehat{\rm PD}=[p]\setminus\widehat{\rm ND}$ contains all of the defective items, and at most p^ξ non-defectives, for some $\xi\in(0,\theta)$ (this will later be taken arbitrarily close to θ). We denote the complements of these events by $P_{\rm e,1}^{\rm (D)}$ and $P_{\rm e,1}^{\rm (ND)}$ respectively.

Analysis of defective items. For any defective item $j\in S$, the number of negative tests in which j is included is distributed as $N_{{\rm neg},j}\sim{\rm Binomial}\big(n,\frac{\rho\nu}{k}\big)$. Moreover,

$$P_{\rm e,1}^{\rm (D)} = \mathbb{P}\Big[\bigcup_{j\in S}\{j\in\widehat{\rm ND}\}\Big] \tag{99}$$

$$\le k\mathbb{P}\Big[N_{{\rm neg},j}\ge\frac{\alpha n\nu}{k}\Big] \tag{100}$$

$$= k\mathbb{P}\Big[N_{{\rm neg},j}\ge\Big(1+\frac{\alpha-\rho}{\rho}\Big)\cdot\mathbb{E}[N_{{\rm neg},j}]\Big], \tag{101}$$

where (100) holds for an arbitrary fixed $j\in S$ by the union bound and the definition of $\widehat{\rm ND}$ (*cf.*, (88)), and (101) follows since $\mathbb{E}[N_{{\rm neg},j}]=\frac{n\rho\nu}{k}$ and $1+\frac{\alpha-\rho}{\rho}=\frac{\alpha}{\rho}$. Applying the binomial concentration bound (20) to (101), we deduce that

$$P_{\rm e,1}^{\rm (D)} \le k\exp\Big(-\frac{n\rho\nu}{k}\cdot D_1(1+\epsilon_1^{\rm (D)})\Big), \tag{102}$$

where $\epsilon_1^{\rm (D)}=\frac{\alpha-\rho}{\rho}$. As a result, we can achieve $P_{\rm e,1}^{\rm (D)}\to 0$ with a number of tests satisfying

$$n\ge\Big(\frac{1}{\rho\nu}\cdot\frac{1}{D_1(\alpha/\rho)}\cdot k\log k\Big)(1+o(1)), \tag{103}$$

since $1+\epsilon_1^{\rm (D)}=\frac{\alpha}{\rho}$.

Analysis of non-defective items. Let $P_{\rm e,1}^{\rm (ND)}(n_{\rm neg})$ be defined similarly to $P_{\rm e,1}^{\rm (ND)}$, but conditioned on $N_{\rm neg}$ taking a given value $n_{\rm neg}$. It suffices to establish that $P_{\rm e,1}^{\rm (ND)}(n_{\rm neg})\to 0$ for all $n_{\rm neg}$ satisfying (97).

Since the test outcomes depend only on the defective items, one can envision the non-defective items as being placed in each test with probability $\frac{\nu}{k}$ *after* the test outcomes have been produced. As a result, given $N_{\rm neg}=n_{\rm neg}$, the number of negative tests $N_{{\rm neg},j}$ including a given item $j\notin S$ is distributed as $(N_{{\rm neg},j}|n_{\rm neg})\sim{\rm Binomial}\big(n_{\rm neg},\frac{\nu}{k}\big)$. Hence, for any $j\notin S$, we have

$$\mathbb{P}[j\notin\widehat{\rm ND}\,|\,n_{\rm neg}] = \mathbb{P}\Big[N_{{\rm neg},j}<\frac{\alpha n\nu}{k}\,\Big|\,n_{\rm neg}\Big] \tag{104}$$

$$\le\mathbb{P}\Big[N_{{\rm neg},j}\le(1-\epsilon_1^{\rm (ND)})\times\mathbb{E}[N_{{\rm neg},j}|n_{\rm neg}](1+o(1))\,\Big|\,n_{\rm neg}\Big], \tag{105}$$

where $\epsilon_1^{\rm (ND)}$ is defined in such a way that $\frac{\alpha n\nu}{k}\le(1-\epsilon_1^{\rm (ND)})\frac{n_{\rm neg}\nu}{k}(1+o(1))$ for all $n_{\rm neg}$ satisfying (97). By some simple re-arrangements, we find that we can choose $\epsilon_1^{\rm (ND)}=1-\frac{\alpha}{e^{-\nu}+\rho(1-e^{-\nu})}$. Applying the concentration bound (19) to (105), we find that

$$\mathbb{P}[j\notin\widehat{\rm ND}\,|\,n_{\rm neg}] \le\exp\Big(\frac{n_{\rm neg}\nu}{k}\cdot D_1(1-\epsilon_1^{\rm (ND)})\cdot(1+o(1))\Big). \tag{106}$$

As a result, letting G denote the number of non-defective items in $\widehat{\rm PD}=[p]\setminus\widehat{\rm ND}$, we find that $\mathbb{E}[G\,|\,n_{\rm neg}]$ is upper bounded by $p-k$ times the right-hand side of (106). Applying Markov's inequality, we obtain

$$\mathbb{P}[G\ge p^\xi\,|\,n_{\rm neg}] \le p^{1-\xi}\exp\Big(\frac{n_{\rm neg}\nu}{k}\cdot D_1(1-\epsilon_1^{\rm (ND)})\cdot(1+o(1))\Big). \tag{107}$$

Since $n_{\rm neg}$ satisfies (97), we find that we can achieve $P_{\rm e,1}^{\rm (ND)}\to 0$ under the condition

$$n\ge\Big(\frac{1-\xi}{\nu}\cdot\frac{1}{e^{-\nu}+\rho(1-e^{-\nu})}\cdot\frac{1}{D_1(\alpha/\zeta)}\cdot k\log p\Big)\times(1+o(1)), \tag{108}$$

since by definition $\epsilon_1^{\rm (ND)}=1-\frac{\alpha}{\zeta}$.

Analysis of Second Step. We follow a similar argument to the one following Lemma 2; due to these similarities, we omit some details and focus on the differences. Note that here we only need to focus on defective items: As long as $\widehat{\rm PD}$ contains all defectives, the the second step will never include a non-defective item in $\widehat{S}$ (*cf.*, (89)), as positive tests must always contain a defective item.

Fix a defective $j\in S$ and consider the triplet $(N_{\rm neg},\widetilde{N}_{{\rm pos},j},N_{\rm other})$, defined similarly to the RZ-channel case:

- $N_{\rm neg}$ is the number of negative tests; for the Z-channel, the associated probability is $q_{\rm neg}=\big(e^{-\nu}+\rho(1-e^{-\nu})\big)(1+o(1))$ (see (96)).
- $\widetilde{N}_{{\rm pos},j}$ is the number of *positive* tests containing j but no other defective item; the probability of a given

test satisfying this condition is $\widetilde{q}_j = (1-\rho)\frac{\nu}{k}\big(1-\frac{\nu}{k}\big)^{k-1} = \big(\frac{(1-\rho)\nu e^{-\nu}}{k}\big)(1+o(1))$.

- $N_{\rm other}$ is the number of remaining tests, with associated probability $q_{\rm other} = 1 - q_{\rm neg} - \widetilde{q}_j$.

Hence, $(N_{\rm neg}, \widetilde{N}_{{\rm pos},j}, N_{\rm other})$ has a multinomial distribution with n trials and parameters $(q_{\rm neg}, \widetilde{q}_j, q_{\rm other})$.

We condition on $(\widehat{\rm pd}, g, n_{\rm neg})$ with $n_{\rm neg}$ satisfying the concentration bound (97), $\widehat{\rm pd}$ containing all of the defective items, and $g \le p^{\xi}$, in accordance with the first step. Similarly to the RZ-channel case, we consider splitting the $\widetilde{N}_{{\rm pos},j}$ tests into sub-classes with counts $(N'_{{\rm pos},j}, \widetilde{N}_{{\rm pos},j} - N'_{{\rm pos},j})$. Similarly to (59), the associated conditional probability of the first sub-class is $\gamma = 1 - o(1)$.

Applying Lemma 2, we deduce that given $(\widehat{\rm pd}, g, n_{\rm neg})$, the triplet $(N'_{{\rm pos},j}, \widetilde{N}_{{\rm pos},j} - N'_{{\rm pos},j}, N_{\rm other})$ is multinomial with $n - n_{\rm neg}$ trials, with the first probability being

$$q'_{{\rm pos},j} = \frac{\widetilde{q}_j \gamma}{1 - q_{\rm neg}} \tag{109}$$

$$= \left(\frac{(1-\rho)\nu e^{-\nu}}{k(1-q_{\rm neg})}\right)(1+o(1)). \tag{110}$$

This gives the following analog of (62):

$$(N'_{{\rm pos},j} \,|\, n_{\rm neg}, \widehat{\rm pd}, g) \sim {\rm Binomial}\bigg(n(1-q_{\rm neg})(1+o(1)), \frac{(1-\rho)\nu e^{-\nu}}{k(1-q_{\rm neg})}(1+o(1))\bigg). \tag{111}$$

Recall from (89) that a given item j is included in the final estimate $\widehat{S}$ if $N'_{{\rm pos},j} > 0$. Consequently, we deduce from (111) that any defective item $j \in S$ yields

$$\mathbb{P}[j \notin \widehat{S} \,|\, n_{\rm neg}, \widehat{\rm pd}, g] = \left(1 - \frac{(1-\rho)\nu e^{-\nu}}{k(1-q_{\rm neg})}(1+o(1))\right)^{n(1-q_{\rm neg})(1+o(1))} \tag{112}$$

$$\le \exp\bigg(-\frac{n}{k}(1-\rho)\nu e^{-\nu}\cdot(1+o(1))\bigg), \tag{113}$$

where we have applied $1-\zeta \le e^{-\zeta}$. By the union bound, the probability of there existing some $j \in S$ failing to be included in $\widehat{S}$ is at most k times the right-hand side of (113). By re-arranging, we deduce that $P_{\rm e,2}^{\rm (D)} \to 0$ as $p \to \infty$ under the condition

$$n \ge \left(\frac{1}{(1-\rho)\nu e^{-\nu}} \cdot k \log k\right)(1+o(1)). \tag{114}$$

Wrapping up. Combining the conditions in (103), (108), and (114), and noting that ξ can be arbitrarily close to θ in (108) (yielding the logarithmic term $(1-\theta)\log p = \big(\log \frac{p}{k}\big)(1+o(1))$), we deduce Theorem 5.

B. Algorithm-Independent Converse

The proof of the converse part of Theorem 2 proceeds along similar lines to the corresponding argument for Theorem 1. We consider the optimal ML decoder of the form (72), identify the suitable error event, analyze the error probability using concentration inequalities, and studying the intersection of relevant D_γ functions.

Let $\widetilde{N}_{{\rm pos},i}$ denote the number of positive tests containing $i \in S$ and no other defectives, and $\widetilde{N}_{{\rm neg},i}$ denote the number of negative tests containing $i \in S$ and no other defectives. The probability of a test containing no defectives is $(1-\nu/k)^k$, so we write $N_{\rm neg}^{\rm (ND)} \sim {\rm Binomial}(n, (1-\nu/k)^k)$ for the number of tests that contain no defectives (rightfully negative). By binomial concentration, we can assume that $N_{\rm neg}^{\rm (ND)} = ne^{-\nu}(1+o(1))$, as this occurs with probability approaching one.

The argument used to prove a converse is the following:

1. We look for a defective item $i \in S$ with $\widetilde{N}_{{\rm pos},i} = 0$ and $\widetilde{N}_{{\rm neg},i} \ge d = \Psi\nu e^{-\nu}\gamma \log p$ (for some Ψ to be determined). We call such a defective item "strongly masked".
2. Given a strongly masked item i, we look for an non-defective item that appears in fewer than d of the $N_{\rm neg}^{\rm (ND)} + \widetilde{N}_{{\rm neg},i}$ of the tests that would be rightfully negative if i were removed from the defective set. We call such an item "weakly intruding".

If there exists such a strongly masked item $i \in S$ and a weakly intruding item $j \notin S$, then since $d \le \widetilde{N}_{{\rm neg},i}$ the ML decoder will prefer the set $S \setminus \{i\} \cup \{j\}$ to the true defective set (assuming $\rho < 1/2$), and hence will make a mistake. We first argue that there exists a strongly masked defective item.

Lemma 5. *Fix $\gamma > 0$, $\rho \le \Psi < 1$, and $d = \Psi\nu e^{-\nu}\gamma\log p = \Psi\nu e^{-\nu} n/k$ with $n = \gamma k \log p$. If $\theta > e^{-\nu}\nu\gamma(D_\rho(\Psi)+1-\rho)$, then with probability tending to one there exists a strongly masked item $i \in S$ (with $\widetilde{N}_{{\rm pos},i} = 0$ and $\widetilde{N}_{{\rm neg},i} \ge d$).*

Proof. For each defective $i \in S$, we write V_i for the indicator of the event that it is strongly masked, and $V = \sum_{i\in S} V_i$ for the total number of strongly masked items. As in [12], we know that $(\widetilde{N}_{{\rm pos},i}, \widetilde{N}_{{\rm neg},i})$ are components of a multinomial distribution, with respective parameters (n, q_+, q_-), where $q_+ = (1-\nu/k)^{k-1}(\nu/k)(1-\rho)$ and $q_- = (1-\nu/k)^{k-1}(\nu/k)\rho$ are the probabilities of a particular test containing defective i and no other defectives, and being positive or negative respectively. Further, observe that

$$(\widetilde{N}_{{\rm neg},i} \,|\, \widetilde{N}_{{\rm pos},i} = 0) \sim {\rm Binomial}(n, q_-/(1-q_+)) \tag{115}$$

by a simpler version of Lemma 2 with $\gamma = 1$.

We will apply the binomial concentration bound (20) and its matching lower bound (23) to $\widetilde{N}_{{\rm neg},i}$ conditioned

on $\widetilde{N}_{\mathrm{pos},i} = 0$. Since we consider the event $\widetilde{N}_{\mathrm{neg},i} \geq d$ and we have assumed $d = \Psi\nu e^{-\nu} n/k$, we introduce a constant ϵ_p defined to satisfy

$$(1+\epsilon_p)\frac{nq_-}{1-q_+} = d = \Psi\nu e^{-\nu}\frac{n}{k}, \tag{116}$$

so that the left hand side is slightly above the conditional mean of $\widetilde{N}_{\mathrm{neg},i}$. Since $kq_+ \to e^{-\nu}\nu(1-\rho)$ and $kq_- \to e^{-\nu}\nu\rho$ by definition, we readily deduce from (116) that $1+\epsilon_p \to \Psi/\rho$ as $p \to \infty$, which is in the range $\big[1, \frac{1}{\rho}\big]$ by the assumption $\Psi \in [\rho, 1]$.

Combining the above and using binomial concentration, we obtain

$$\begin{aligned}
\mathbb{E}[V_i] &= \mathbb{P}[\widetilde{N}_{\mathrm{pos},i} = 0, \widetilde{N}_{\mathrm{neg},i} \geq d] && (117)\\
&= \mathbb{P}[\widetilde{N}_{\mathrm{pos},i} = 0]\mathbb{P}[\widetilde{N}_{\mathrm{neg},i} \geq d | \widetilde{N}_{\mathrm{pos},i} = 0] && (118)\\
&= (1-q_+)^n\, \mathbb{P}\Big[\mathrm{Binomial}\big(n, q_-/(1-q_+)\big) \geq d\Big] && (119)\\
&\leq (1-q_+)^n \exp\left(-\frac{nq_-}{1-q_+}D_1(1+\epsilon_p)\right) && (120)\\
&\leq \exp\left(-n\Big(q_+ + \frac{q_-}{1-q_+}D_1(1+\epsilon_p)\Big)\right) && (121)\\
&= \exp\Big(-\frac{n}{k}(e^{-\nu}\nu)\,(1-\rho+\rho D_1(\Psi/\rho)) \\
&\qquad\qquad \times (1+o(1))\Big) && (122)\\
&= p^{-\gamma e^{-\nu}\nu(1-\rho+D_\rho(\Psi))\cdot(1+o(1))}, && (123)
\end{aligned}$$

where (119) follows from (115), (120) follows from the concentration bound (20),(121) uses $1-q_+ \leq e^{-q_+}$, (122) applies the above-mentioned asymptotics of (q_+, q_-, ϵ_p), and (123) follows since $n = \gamma k \log p$.

We can apply the same argument to lower bound $\mathbb{E}[V_i]$ via (23). As in the proof of Lemma 4 above the sharpness factor is at least $p^{-\epsilon'}$ for any $\epsilon' > 0$ and p sufficiently large; in other words, it behaves as $p^{o(1)}$. Combining this lower bound with the upper bound (123), we conclude that $\mathbb{E}[V_i] = p^{-\gamma e^{-\nu}\nu(1-\rho+D_\rho(\Psi))\cdot(1+o(1))+o(1)}$. Hence, since $k \asymp p^\theta$, the expectation $\mathbb{E}[V] = k\mathbb{E}[V_i]$ tends to infinity under the assumed condition $\theta > \gamma e^{-\nu}\nu(1-\rho+D_\rho(\Psi))$.

Furthermore, we can bound the variance of V by a negative association argument (using [45, Property P3] as above) to deduce that for any two defectives $i \neq j$:

$$\begin{aligned}
&\mathbb{E}[V_iV_j] \\
&= \mathbb{P}[\widetilde{N}_{\mathrm{pos},i} = 0, \widetilde{N}_{\mathrm{neg},i} \geq d, \widetilde{N}_{\mathrm{pos},j} = 0, \widetilde{N}_{\mathrm{neg},j} \geq d] && (124)\\
&\leq (1-q_+)^{2n}\left(\mathbb{P}[\mathrm{Binomial}(n, q_-/(1-2q_+)) \geq d]\right)^2. && (125)
\end{aligned}$$

Here the factor of $1-2q_+$ can by understood as multinomial conditioning (*cf.*, Lemma 2) on *both* $\widetilde{N}_{\mathrm{pos},i} = 0$ and $\widetilde{N}_{\mathrm{pos},j} = 0$. Continuing, we write

$$\begin{aligned}
\frac{\mathrm{Var}(V)}{(\mathbb{E}[V])^2} &= \frac{k\mathrm{Var}[V_i] + k(k-1)\mathrm{Cov}[V_i, V_j]}{k^2(\mathbb{E}[V_i])^2} && (126)\\
&= \frac{1}{k\mathbb{E}[V_i]} + \frac{\mathbb{E}[V_iV_j]}{(\mathbb{E}[V_i])^2} - 1, && (127)
\end{aligned}$$

where we used $\mathrm{Var}[V_i] = \mathbb{E}[V_i] - (\mathbb{E}[V_i])^2$ for binomial random variables, and the fact that $\mathrm{Cov}[V_i, V_j] = \mathbb{E}[V_iV_j] - \mathbb{E}[V_i]\mathbb{E}[V_j] = \mathbb{E}[V_iV_j] - (\mathbb{E}[V_i])^2$. We have already established that $k\mathbb{E}[V_i] \to \infty$, so the first term in (127) tends to zero as $p \to \infty$. In addition, from (119) and (125), we have

$$\begin{aligned}
\frac{\mathbb{E}[V_iV_j]}{(\mathbb{E}[V_i])^2} &\leq \left(\frac{\mathbb{P}\big[\mathrm{Binomial}\big(n, \frac{q_-}{1-2q_+}\big) \geq d\big]}{\mathbb{P}\big[\mathrm{Binomial}\big(n, \frac{q_-}{1-q_+}\big) \geq d\big]}\right)^2 && (128)\\
&\leq 1 + o(1), && (129)
\end{aligned}$$

where (129) is proved in Appendix F.

Combining the preceding observations with (127) gives $\frac{\mathrm{Var}(V)}{(\mathbb{E}[V])^2} \to 0$. This implies that V is sufficiently concentrated around its mean to deduce (via Chebyshev's inequality) that there exists a strongly masked defective (i.e., $V > 0$) with probability approaching one, as desired. □

We now argue that there exists a weakly intruding non-defective. First note that the analysis so far has only considered the columns of the test matrix corresponding to defective items, and the i.i.d. design of the test matrix means that columns corresponding to non-defectives are independent of those. Recall that an item is weakly intruding if it appears in fewer than d of the $N_{\mathrm{neg}}^{(\mathrm{ND})} + \widetilde{N}_{\mathrm{neg},i}$ tests that would be rightfully negative if i were removed, and that $N_{\mathrm{neg}}^{(\mathrm{ND})} = ne^{-\nu}(1+o(1))$ (*cf.*, start of this subsection). Again using the binomial concentration bound in (20) along with $\widetilde{N}_{\mathrm{neg},i} \sim \mathrm{Binomial}(n, q_-/(1-q_+))$ (*cf.*, proof of Lemma 5), we can also assume that $\widetilde{N}_{\mathrm{neg},i} \leq Ce^{-\nu}\nu n/k$ for fixed $C > \rho$, as this holds with probability approaching one.

Lemma 6. *Fixing $\gamma > 0$, $\Psi < 1$, and $C > \rho$, and setting $d = \Psi\nu e^{-\nu}\gamma\log p$ and $n = \gamma k \log p$, conditioned on there existing a strongly masked $i \in S$ with $\widetilde{N}_{\mathrm{neg},i} \leq Ce^{-\nu}\nu n/k$, there also exists a weakly intruding non-defective item with probability tending to one if $1 > \gamma\nu e^{-\nu}D_1(\Psi)$.*

Proof. Conditioned on $N_{\mathrm{neg}}^{(\mathrm{ND})}$ and $\widetilde{N}_{\mathrm{neg},i}$, the number of rightfully negative tests that each non-defective item appears in is distributed as $\mathrm{Binomial}(N_{\mathrm{neg}}^{(\mathrm{ND})} + \widetilde{N}_{\mathrm{neg},i}, \nu/k)$ (note that the relevant columns of $\mathbf{X}$ for non-defective items are independent of those that determine $N_{\mathrm{neg}}^{(\mathrm{ND})}$ and $\widetilde{N}_{\mathrm{neg},i}$). By the above-mentioned assumptions $N_{\mathrm{neg}}^{(\mathrm{ND})} = ne^{-\nu}(1+o(1))$ and $\widetilde{N}_{\mathrm{neg},i} \leq Ce^{-\nu}\nu n/k$, the probability of this binomial random variable being less than d is

lower bounded by that of $Z \sim \text{Binomial}(ne^{-\nu}(1+\delta_p + C/k), \nu/k)$, where $\delta_p = o(1)$.

Defining ϵ'_p to be such that

$$(1-\epsilon'_p)\frac{n}{k}e^{-\nu}\nu(1+\delta_p+C/k) = d = \Psi\nu e^{-\nu}\frac{n}{k}, \quad (130)$$

we observe that $1-\epsilon'_p \to \Psi$. Hence, we have $(1+\delta_p+C/k)D_1(1-\epsilon'_p) = D_1(\Psi)(1+o(1))$. In addition, when bounding $\mathbb{P}[Z<d]$, the discussion following (23) reveals that the concentration bound (19) has a matching lower bound with a $p^{-\epsilon'}$ pre-factor. Combining the above observations, we obtain for any $j \notin S$ that

$$\begin{aligned}
&\mathbb{P}[\text{ item } j \text{ is weakly intruding }] \\
&\geq \mathbb{P}[Z<d] && (131)\\
&\geq p^{-\epsilon'}\exp\left(-\frac{n}{k}e^{-\nu}\nu(1+\delta_p+C/k)D_1(1-\epsilon'_p)\right) && (132)\\
&\geq p^{-\epsilon'}\exp\left(-\gamma(\log p)e^{-\nu}\nu D_1(\Psi)\cdot(1+o(1))\right), && (133)
\end{aligned}$$

where and (133) uses $n = \gamma k \log p$ in addition to the above observations.

By (133), the expected number of weakly intruding non-defective items is at least

$$(p-k)\cdot p^{-\epsilon'}\cdot p^{-\gamma\nu e^{-\nu}D_1(\Psi)\cdot(1+o(1))},$$

which grows to infinity as $\Omega(p^\tau)$ for some $\tau > 0$, due to the assumption $1 > \gamma\nu e^{-\nu}D_1(\Psi)$ and the fact that ϵ' may be arbitrarily small. Finally, the i.i.d. design of the matrix means that non-defective items are weakly intruding independently of one another, so by binomial concentration, the actual number of of weakly intruding items is positive with probability approaching one. □

By the definition of rate in (4) along with $n = \gamma k \log p$ and $k \asymp p^\theta$, we can rephrase Lemma 5 to say that there exists a strongly masked defective with high probability when the rate $R \sim \frac{1-\theta}{\gamma \log 2}$ satisfies

$$\frac{1-\theta}{\gamma\log 2} > \left(\frac{(1-\theta)\nu e^{-\nu}}{\log 2}\right)\frac{D_\rho(\Psi)+1-\rho}{\theta}. \quad (134)$$

Similarly, Lemma 6 states that there exists a weakly intruding item with high probability if

$$\frac{1-\theta}{\gamma\log 2} > \left(\frac{(1-\theta)\nu e^{-\nu}}{\log 2}\right)D_1(\Psi). \quad (135)$$

Combining (134) and (135), we see that the error event will occur if the rate is bigger than

$$\frac{(1-\theta)\nu e^{-\nu}}{\log 2}\max\left\{\frac{D_\rho(\Psi)+1-\rho}{\theta}, D_1(\Psi)\right\}. \quad (136)$$

We optimize this expression with respect to Ψ in Appendix E to complete the proof of Theorem 2.

V. Conclusion

We have introduced and analyzed variants of the definite defectives (DD) algorithm for noisy group testing, with an emphasis on the Z-channel and reverse Z-channel models. Under RZ noise, our achievability result (part of which also uses the COMP algorithm) matches an algorithm-independent converse for Bernoulli testing for a broad range of dense scaling regimes, and matches a converse specific to DD for Bernoulli testing in sparse regimes. While more significant gaps remain for the Z-channel and general binary channel (see Appendix I), the bounds are still matching or near-matching in several low-noise high-sparsity regimes. Further closing these gaps, either by improved achievability or improved converse bounds (or both), poses an interesting direction for further research.

Appendix

A. Technical Lemma Regarding Intersection of D_γ Functions

We present a result giving an explicit formula for the intersection of the D_γ functions introduced in (14), i.e., $D_\gamma(t) = t\log\frac{t}{\gamma} - t + \gamma$.

Lemma 7. *Fixing $0 \leq \gamma_1 \leq \gamma_2$, $c \geq 0$, and $d \geq 0$, we have the following:*

1. *The equation*

$$\varphi(t) := D_{\gamma_1}(t) - cD_{\gamma_2}(t) + d = 0 \quad (137)$$

has a unique solution for $t \in [\gamma_1, \gamma_2]$ if and only if

$$d \leq cD_{\gamma_2}(\gamma_1). \quad (138)$$

2. *If* (138) *fails, the smallest value of $\varphi(t)$ is achieved by $t = \gamma_1$, and equals $-cD_{\gamma_2}(\gamma_1) + d > 0$.*
3. *If* (138) *holds, the solution to* (137) *is given by*

$$\begin{aligned}
t^* &= \frac{\gamma_2-\gamma_1-d}{\log(\gamma_2/\gamma_1)} \quad \text{for } c=1 && (139)\\
t^* &= \gamma_1^{1/(1-c)}\gamma_2^{-c/(1-c)}e^{z^*+1} \quad \text{for } c \neq 1 && (140)\\
&= -\frac{\gamma_1+d-c\gamma_2}{(1-c)z^*}, && (141)
\end{aligned}$$

where z^* is a solution to

$$ze^z = -\frac{1}{e}\frac{1}{1-c}\left(1+\frac{d-c\gamma_2}{\gamma_1}\right)\left(\frac{\gamma_2}{\gamma_1}\right)^{c/(1-c)}. \quad (142)$$

This can be found using the appropriate branch of the Lambert W-function.

Proof. We have the following:

1. Observe that $D_{\gamma_1}(t)$ is strictly increasing for $t \in (\gamma_1, \gamma_2)$ and $D_{\gamma_2}(t)$ is strictly decreasing for t in this range, so $\varphi(t)$ is strictly increasing in $t \in (\gamma_1, \gamma_2)$. Furthermore, $\varphi(\gamma_2) = D_{\gamma_1}(\gamma_2) + d \geq 0$ by definition, and $\varphi(\gamma_1) = -cD_{\gamma_2}(\gamma_1) + d$, which is non-positive

if and only if (138) holds. In other words, (138) is equivalent to the existence of a change of sign of φ at some point in $[\gamma_1, \gamma_2]$.

2. If (138) fails then $\varphi(\gamma_1) > 0$, so since φ is increasing we know that it is minimized at this point.
3. For $c = 1$, we can solve directly using the fact that

$$\begin{aligned}\varphi(t) &= D_{\gamma_1}(t) - D_{\gamma_2}(t) + d \\ &= t \log\left(\frac{\gamma_2}{\gamma_1}\right) + \gamma_1 - \gamma_2 + d, \end{aligned} \tag{143}$$

yielding (139). For $c \neq 1$, we consider reparameterizing $t = \gamma_1^{1/(1-c)}\gamma_2^{-c/(1-c)}e^{z+1}$ by another variable z to obtain

$$D_{\gamma_1}(t) = t \log\left(\frac{\gamma_1^{1/(1-c)}\gamma_2^{-c/(1-c)}e^{z}}{\gamma_1}\right) + \gamma_1 \tag{144}$$

$$D_{\gamma_2}(t) = t \log\left(\frac{\gamma_1^{1/(1-c)}\gamma_2^{-c/(1-c)}e^{z}}{\gamma_2}\right) + \gamma_2. \tag{145}$$

Since $\frac{1}{1-c} - 1 = \frac{c}{1-c}$ and $-\frac{c}{1-c} - 1 = -\frac{1}{1-c}$, we find that evaluating $D_{\gamma_1}(t) - cD_{\gamma_2}(t)$ leads to a cancellation of powers of γ_1 and γ_2 inside the logarithm, and hence

$$\varphi(t) = D_{\gamma_1}(t) - cD_{\gamma_2}(t) + d \tag{146}$$

$$= tz(1-c) + \gamma_1 - c\gamma_2 + d \tag{147}$$

$$= e(1-c)\gamma_1^{1/(1-c)}\gamma_2^{-c/(1-c)} \cdot ze^{z} + (d + \gamma_1 - c\gamma_2), \tag{148}$$

and (142) follows. We deduce (140) from the definition of t, and (141) by equating (147) with zero.

□

B. Proving Theorem 1 (achievability part, RZ noise) via Theorem 3

We consider the rates achievable by COMP and NDD separately (*cf.*, Lemma 1 and Theorem 3). Since $\frac{k\log_2(p/k)}{k\log p} \sim \frac{1-\theta}{\log 2}$, Lemma 1 tells us that the limiting rate in (4) for COMP becomes

$$\begin{aligned} R &\sim \frac{k\log_2\frac{p}{k}}{n_{\mathrm{COMP}}} = \frac{k\log_2\frac{p}{k}}{k\log p}\frac{k\log p}{n_{\mathrm{COMP}}} \\ &\sim \frac{(1-\theta)}{\log 2}(1-\rho)\nu e^{-\nu} = \frac{(1-\theta)(1-\rho)}{e\log 2}, \end{aligned} \tag{149}$$

taking the choice of $\nu = 1$ that maximizes $\nu e^{-\nu}$.

To find the rate achievable by NDD, we rewrite (43), (44) and (45) respectively, using (15) in the third case, to obtain

$$\frac{k\log p}{n_1^{(D)}} = \nu e^{-\nu}\frac{1-\rho}{1-\xi} \tag{150}$$

$$\frac{k\log p}{n_2^{(D)}} = \nu e^{-\nu}\frac{D_1(\beta)}{\theta} \tag{151}$$

$$\frac{k\log p}{n_2^{(ND)}} = \nu e^{-\nu}\frac{\rho D_1(\beta/\rho)}{\xi} = \nu e^{-\nu}\frac{D_\rho(\beta)}{\xi}, \tag{152}$$

since $k\log_2(p/k) = \big((1-\theta)k\log_2 p\big)(1+o(1))$.

Note that (150) is increasing in ξ, (152) is decreasing in ξ, (151) is decreasing in $\beta \in (\rho, 1)$ and (152) is increasing in $\beta \in (\rho, 1)$, so we need to choose the parameters ξ and β to balance these terms. Using a similar argument to (149) above, and again making the optimal choice $\nu = 1$, the limiting rate in (4) becomes

$$\begin{aligned} R &\sim \frac{k\log_2(p/k)}{n} \\ &\sim \frac{(1-\theta)}{e\log 2}\min\left\{\frac{1-\rho}{1-\xi}, \frac{D_1(\beta)}{\theta}, \frac{D_\rho(\beta)}{\xi}\right\}. \end{aligned} \tag{153}$$

Since D_1 and D_ρ in (153) are continuous in their arguments, we can consider choosing $\beta \in [\rho, 1]$ and $\xi \in [0, \theta]$,[4] where previously we excluded the endpoints. If there exists a value $\beta^* \in [\rho, 1]$ that makes

$$\frac{D_1(\beta^*)}{\theta} = D_\rho(\beta^*) + 1 - \rho, \tag{154}$$

then taking $\xi = \xi^* := \frac{D_\rho(\beta^*)}{D_\rho(\beta^*)+1-\rho}$ (if valid, i.e., in $[0, \theta]$) would make all bracketed terms in (153) become equal to $D_\rho(\beta^*) + 1 - \rho$, suggesting that a putative rate of

$$\frac{(1-\theta)(D_\rho(\beta^*) + 1 - \rho)}{e\log 2} \tag{155}$$

might be possible. We can consider whether there exists a solution to (154) by taking $\gamma_1 = \rho$, $\gamma_2 = 1$, $c = \frac{1}{\theta}$ and $d = 1 - \rho$ in Lemma 7. Examining (138), there exists a solution $\beta^* \in (\rho, 1)$ if and only if $(1-\rho)\theta \leq D_1(\rho) = \rho\log\rho + 1 - \rho$, or equivalently if $\theta \leq \theta_{\mathrm{crit}}^{(\mathrm{RZ})} = 1 + \frac{\rho\log\rho}{1-\rho}$ (*cf.*, (25)).

Case 1 ($\theta \leq \theta_{\mathrm{crit}}^{(\mathrm{RZ})}$). Equation (140) in Lemma 7 shows that $\beta^* = \rho^{1/(1-1/\theta)}e^{z+1} = \rho^{\theta/(\theta-1)}e^{z+1}$, where z is a solution to $ze^z = -e^{-1}\rho^{\theta/(1-\theta)}$ (see (142), and observe that $\frac{1}{1-c}\big(1 + \frac{d-c\gamma_2}{\gamma_1}\big) = \frac{1}{1-1/\theta}\big(1 + \frac{1-\rho-1/\theta}{\rho}\big) = \frac{1}{\rho}$), i.e., $z = -\kappa$ in the notation of (24). Substituting the value of

[4] As $\xi \to 0$, the third term in (153) grows unbounded, so for $\xi = 0$ we simply lower bound the minimum of three terms by that of the first two terms.

β^* into the definition of D_ρ and applying some algebra,[5] we obtain

$$(1-\theta)(D_\rho(\beta^*)+1-\rho) = \frac{-\log\rho}{\kappa(\theta)}. \tag{156}$$

We need to verify whether the corresponding parameter ξ^* satisfies $\xi^* < \theta$, which is equivalent to $1-\xi^* > 1-\theta$ and in turn (recalling $\xi^* = \frac{D_\rho(\beta^*)}{D_\rho(\beta^*)+1-\rho}$) to

$$(1-\theta)(D_\rho(\beta^*)+1-\rho) < 1-\rho, \tag{157}$$

or equivalently $-\log\rho < \kappa(1-\rho)$ (*cf.*, (156)). Direct calculation shows that this is satisfied if and only if $\theta > \theta_{\rm opt}$, where $\theta_{\rm opt}$ is defined in (26); this is deduced by substituting the "endpoint" value $\kappa = \frac{-\log\rho}{1-\rho}$ into $\kappa e^{-\kappa} = -e^{-1}\rho^{\theta/(1-\theta)}$.[6]

Hence for $\theta_{\rm opt} < \theta \le \theta_{\rm crit}^{\rm (RZ)}$, using the fact that $(1-\theta)(D_\rho(\beta^*)+1-\rho) = \frac{-\log\rho}{\kappa(\theta)}$ as per (155), we obtain a rate of $\frac{-\log\rho}{\kappa(\theta)e\log 2}$ as claimed.

For $\theta \le \theta_{\rm opt}$, ξ^* is not a legitimate choice, but we can obtain a rate of $\frac{1-\rho}{e\log 2}$ by picking $\beta = \beta^*$ and $\xi = \theta$. In this case, the first term (i.e., $\frac{1-\rho}{1-\theta}$) provides the minimum in (153) since the reverse of (157) holds, implying

$$\frac{1-\rho}{1-\theta} \le D_\rho(\beta^*)+1-\rho = \frac{D_1(\beta^*)}{\theta}, \tag{158}$$

$$\frac{1-\rho}{1-\theta} \le \frac{D_\rho(\beta^*)}{\theta}, \tag{159}$$

where the equality in the first expression applies (154), and the second expression follows by rewriting the reverse of (157) as $(1-\theta)D_\rho(\beta^*) \ge (1-\rho)-(1-\theta)(1-\rho) = \theta(1-\rho)$.

Case 2 ($\theta > \theta_{\rm crit}^{\rm (RZ)}$). In this case, the second part of Lemma 7 tells us that the optimal choice is to take $\beta = \rho$ and ξ arbitrarily close to zero (to keep the third term in (153) zero while maximizing the first term). However, in analogy with (154), for $\theta > \theta_{\rm crit}^{\rm (RZ)}$ it holds that $\frac{D_1(\rho)}{\theta} < 1-\rho$, so the minimum in (153) is strictly smaller than the first term. In other words, the optimized NDD rate is strictly less than $\frac{(1-\theta)(1-\rho)}{e\log 2}$, which we know from (149) above is achievable by COMP. In other words, for sufficiently dense problems, the NDD rate bound of Theorem 3 is worse than that attained by COMP. Therefore, in this regime, we get the required rate in Theorem 3 from COMP instead of NDD.

[5]Write $ze^z = -e^{-1}\rho^{\theta/(1-\theta)}$ as $e^{z+1}\rho^{\theta/(\theta-1)} = -\frac{1}{z}$, or equivalently $\beta^* = \frac{1}{\kappa}$. Then $D_\rho(\beta^*)+1-\rho = \beta^*\log\frac{\beta^*}{\rho} - \beta^* + 1 = \beta^*\log\frac{\beta^*}{\rho e} + 1 = \beta^*\frac{-\log\rho}{1-\theta} - \beta^*\kappa + 1 = \beta^*\frac{-\log\rho}{1-\theta}$, since $\frac{\beta^*}{\rho e} = \rho^{-1/(1-\theta)}e^{-\kappa}$.

[6]In more detail, this choice gives $\log\left(\kappa e^{-\kappa}\right) = \frac{\log\rho}{1-\rho} + \log\left(\frac{-\log\rho}{1-\rho}\right)$, and we can also rewrite $\kappa e^{-\kappa} = -e^{-1}\rho^{\theta/(1-\theta)}$ as $\theta = \left(1-\frac{\log\rho}{1+\log(\kappa e^{-\kappa})}\right)^{-1} = \frac{1+\log\left(\kappa e^{-\kappa}\right)}{1+\log\left(\kappa e^{-\kappa}\right)-\log\rho}$ by applying $\frac{\theta}{1-\theta} = \frac{1}{1/\theta-1}$ and re-arranging. Combining these two facts gives $\theta = \theta_{\rm opt}$.

C. Proving Theorem 1 (converse part, RZ noise) via (87)

By the discussion following (87), we want to find the value of Φ that gives the smallest value of

$$\max\left\{\frac{D_1(\Phi)}{\theta}, D_\rho(\Phi)+1-\rho\right\}. \tag{160}$$

This is precisely the problem considered in (154) (without any need to consider constraints on ξ) and recall the following observations that we established via Lemma 7.

Case 1 ($\theta \le \theta_{\rm crit}^{\rm (RZ)}$). We know that the smallest value is given by the intersection of the two curves. In accordance with (156), if

$$R > \frac{-\log\rho}{\kappa(\theta)e\log 2} \tag{161}$$

then the success probability of the ML algorithm tends to zero (again making the optimal choice $\nu = 1$)

Case 2 ($\theta \ge \theta_{\rm crit}^{\rm (RZ)}$). Recall from the arguments following (154) that the maximum of $\frac{D_1(\Phi)}{\theta}$ and $D_\rho(\Phi)+1-\rho$ is achieved by $D_\rho(\Phi)+1-\rho$, and the smallest such value is attained when $\Phi = \rho$. Therefore, if the rate satisfies

$$R \ge \left(\frac{1-\theta}{e\log 2}\right)(1-\rho), \tag{162}$$

then the success probability of the ML algorithm tends to zero (again using the fact that $\nu = 1$ maximizes $\nu e^{-\nu}$).

Finally, the presence of the (reverse) Z-channel capacity $C_{\rm Z}(\rho)$ in Theorem 1 needs no further justification, as such a bound was proved in [11].

D. Proving Theorem 2 (achievability part, Z noise) via Theorem 4

In the following, recall that $\zeta = e^{-\nu}+\rho(1-e^{-\nu}) \ge \rho$. In a similar way to the proof of Theorem 1, we can rewrite (91), (92) and (93) as

$$\frac{k\log k}{n_1^{(D)}} = \nu D_\rho(\alpha) \tag{163}$$

$$\frac{k\log\frac{p}{k}}{n_1^{(ND)}} = \nu D_\zeta(\alpha) \tag{164}$$

$$\frac{k\log k}{n_2^{(D)}} = \nu e^{-\nu}(1-\rho). \tag{165}$$

Using the fact that $\frac{k\log_2(p/k)}{k\log k} = \frac{1-\theta}{\theta\log 2}(1+o(1))$, we deduce that for any choice of parameters ν and α, an achievable rate is given by

$$\frac{1}{\log 2}\min\left\{\frac{(1-\theta)\nu D_\rho(\alpha)}{\theta}, \nu D_\zeta(\alpha), \frac{(1-\theta)(1-\rho)\nu e^{-\nu}}{\theta}\right\}. \tag{166}$$

Again, we maximize (166), first equating the first two terms using Lemma 7 with $\gamma_1 = \rho$, $\gamma_2 = \zeta$, $c = \frac{\theta}{1-\theta}$ and $d = 0$, for which (138) trivially holds and so a unique

α solving $\frac{(1-\theta)\nu D_\rho(\alpha)}{\theta} = \nu D_\zeta(\alpha)$ exists. If $\theta = \frac{1}{2}$, then (139) gives

$$\alpha^*(1/2) = \frac{\zeta - \rho}{\log(\zeta/\rho)} = \frac{\rho s}{\log(1+s)},$$

since $s = \frac{\zeta}{\rho} - 1$ by the definitions of s and ζ. Note that the required[7] bound $\rho \le \alpha^*(1/2) \le \zeta$ holds because it is equivalent to $\log(1+s) \le s \le (1+s)\log(1+s)$. We obtain (32b) by substituting this value in $D_\rho(\alpha)$ in (166) (note that $D_\rho(\alpha^*(1/2)) = \frac{\rho s}{\log(1+s)} \cdot \log \frac{s}{\log(1+s)} - \frac{\rho s}{\log(1+s)} + \rho$ and $\rho s = \zeta - \rho = (1-\rho)e^{-\nu}$).

In the case $\theta \neq \frac{1}{2}$, (142) can be expressed in the form[8]

$$e^{z^*} z^* = -\frac{1}{e} g(s,\theta), \tag{167}$$

for the function g given in (29). This is solved by taking $z^* = \lambda(\theta)$ with $\lambda(\theta)$ given in (30). Our choice of branch of the Lambert W-function is justified in Remark 4 below. The value of α^* given in (31) then follows via (141).[9]

Next, note the following two facts:

- By (140), we have $\frac{\alpha^*}{\rho e} = \left(\frac{1}{1+s}\right)^{c/(1-c)} e^{z^*}$ (this is established via $\rho^{\frac{1}{1-c}} = \rho \cdot \rho^{\frac{c}{1-c}}$ and $\frac{\rho}{\zeta} = \frac{1}{1+s}$).
- By (141), we have $\alpha^* z^* = -\rho - \frac{\rho s \theta}{2\theta - 1}$ (this is established via $-\frac{\rho - \frac{\theta}{1-\theta}\zeta}{1 - \frac{\theta}{1-\theta}} = -\frac{\rho(1-2\theta) - (\zeta - \rho)}{1-2\theta}$ and $\zeta - \rho = \rho s$).

Using the definition of D_ρ followed by these two facts, the first (and therefore also the second) bracketed term in (166) equates to

$$\begin{aligned}
&\frac{(1-\theta)\nu D_\rho(\alpha^*)}{\theta} \\
&= \frac{(1-\theta)\nu}{\theta}\left(\alpha^* \log\left(\frac{\alpha^*}{\rho e}\right) + \rho\right) && (168)\\
&= \frac{(1-\theta)\nu}{\theta}\left((\alpha^* z^* + \rho) - \alpha^* \frac{c}{1-c}\log(1+s)\right) && (169)\\
&= \frac{(1-\theta)\nu}{\theta}\left(-\frac{\rho s \theta}{2\theta - 1} + \alpha^* \frac{\theta}{2\theta-1}\log(1+s)\right) && (170)\\
&= \frac{(1-\theta)(1-\rho)\nu e^{-\nu}}{2\theta - 1}\left(-1 + \frac{\alpha^* \log(1+s)}{\rho s}\right), && (171)
\end{aligned}$$

where (171) uses $\rho s = \zeta - \rho = (1-\rho)e^{-\nu}$. We obtain the desired rate in (32a) upon substituting into (166) and noting that $\frac{\log(1+s)}{\rho s} = \alpha^*(1/2)^{-1}$.

[7]We may again include the endpoints $\alpha = \rho$ and $\alpha = \zeta$ due to the continuity of D_ρ and D_ζ, similarly to the reverse Z-channel model.

[8]The relevant terms are evaluated as follows: (i) $\frac{1}{1-c} = \frac{1-\theta}{1-2\theta}$, (ii) $\left(1 + \frac{d - c\gamma_2}{\gamma_1}\right) = 1 - \frac{\theta}{1-\theta}\frac{\zeta}{\rho}$, (iii) $\left(\frac{\gamma_2}{\gamma_1}\right)^{c/(1-c)} = \left(\frac{\zeta}{\rho}\right)^{\frac{\theta}{1-2\theta}}$. The latter two of these are further simplified using $\frac{\zeta}{\rho} = 1 + s$.

[9]In more detail, $-\frac{\gamma_1 + d - c\gamma_2}{(1-c)z^*}$ evaluates to $-\frac{\rho - \frac{\theta}{1-\theta}\zeta}{(1 - \frac{\theta}{1-\theta})\lambda}$. Multiplying and dividing by ρ and using $\frac{\zeta}{\rho} = 1 + s$, this simplifies to $-\frac{\rho}{\lambda}\left(\frac{1 - \theta - \theta(1+s)}{1-2\theta}\right)$ and in turn to (31).

Remark 4. The choice of branch of the Lambert W-function in (30) follows from the fact that (using the parametrization $t = \gamma_1^{1/(1-c)}\gamma_2^{-c/(1-c)}e^{z+1}$) we require $\rho \le \rho(1+s)^{\theta/(2\theta-1)}e^{z+1} \le \zeta$ in Lemma 7. Rearranging, and using $\frac{\zeta}{\rho} = 1 + s$, the lower bound is equivalent to the fact that

$$z \ge -1 - \frac{\theta}{2\theta - 1}\log(1+s), \tag{172}$$

and the upper bound is equivalent to the fact that

$$z \le -1 - \frac{1-\theta}{2\theta - 1}\log(1+s). \tag{173}$$

Hence, for $\theta < \frac{1}{2}$, (172) tells us that $z \ge -1$, so we need to take the W_0 branch. Similarly, for $\theta > \frac{1}{2}$, (173) tells us that $z \le -1$, so we take the W_{-1} branch.

We can now justify in more detail some of the claims made earlier in Remark 3 regarding the behavior of the achievable rate.

Remark 5. 1. For given s, the function $\alpha^*(\theta)$ is continuous at $\theta = \frac{1}{2}$. This follows because as $\theta \to \frac{1}{2}$ from below, $-e^{-1}g(s,\theta) \to \infty$, and (17) gives $W_0(x) \sim \log x$ for x large, yielding

$$\begin{aligned}
&\lambda(\theta) \sim \log(-e^{-1}g) \\
&= -\frac{\theta}{2\theta - 1}\log(1+s) - 1 + \log\left(\frac{\theta s}{1 - 2\theta} - 1\right)
\end{aligned} \tag{174}$$

by the definition of g. (Here $\sim$ means that the ratio of the terms tends to 1). Similarly, as $\theta \to \frac{1}{2}$ from above, $-e^{-1}g(s,\theta) \to 0$, and (18) gives $W_{-1}(x) \sim \log(-x)$ for x close to zero, yielding

$$\begin{aligned}
&\lambda(\theta) \sim \log(e^{-1}g) \\
&= -\frac{\theta}{2\theta - 1}\log(1+s) - 1 + \log\left(\frac{\theta s}{2\theta - 1} + 1\right).
\end{aligned} \tag{175}$$

In either case, we deduce that $(2\theta - 1)\lambda(\theta) \to -\frac{\log(1+s)}{2}$, and $\alpha^*(\theta) \to \frac{\rho s}{\log(1+s)}$ so the definition in (31) is continuous at this point. Since $\alpha^*(\theta)$ is continuous at $\theta = 1/2$, the definition of the rate $\underline{\mathrm{R}}^{(\mathrm{Z})}(\theta,\rho)$ given in (32) is also continuous at $\theta = 1/2$ (since $\underline{\mathrm{R}}^{(\mathrm{Z})}$ is obtained by substituting α^* into a continuous function $D_\rho(\alpha)$). By (17), we know that $W_0(x) \le \log x$ for x sufficiently large, and hence $\alpha^*(\theta) < \alpha^*(1/2)$ for θ in some left-neighborhood of $1/2$. Similarly, $W_1(x) \le \log(-x)$ for all x sufficiently large (and negative), and hence $\alpha^*(\theta) > \alpha^*(1/2)$ for θ in some right-neighborhood of $1/2$.

2. In the limit as ρ tends to zero, a sub-optimal but useful choice of α is

$$\alpha = \frac{e^{-\nu}\theta}{1-\theta}\frac{1}{(-\log\rho)}, \tag{176}$$

which is greater than ρ (as required for $\alpha > \rho$) when ρ is sufficiently small. Using $aD_b(t) = D_{ab}(at)$ (a simple generalization of (15)), the first two bracketed terms of (166) simplify to

$$\frac{\nu e^{-\nu}}{\log 2} \min\left\{ \frac{(1-\theta)e^{\nu}}{\theta} D_{\rho}(\alpha), e^{\nu} D_{\zeta}(\alpha) \right\} = \frac{\nu e^{-\nu}}{\log 2} \min\left\{ D_{\rho(1-\theta)e^{\nu}/\theta}\left(\frac{\alpha(1-\theta)e^{\nu}}{\theta}\right), D_{e^{\nu}\zeta}(\alpha e^{\nu}) \right\}$$

and under the above choice of α, both $D_{(\cdot)}$ terms tend to 1 as $\rho \to 0$.[10] Setting $\nu = 1$, we obtain $\frac{1}{e\log 2}$, and taking into account the third term of (166) we recover the noiseless rate of (5) in the limit as $\rho \to 0$.

E. Proving Theorem 2 (converse part, Z noise) via (136)

Recall from (136) that the error probability cannot vanish when the rate is above

$$\frac{(1-\theta)\nu e^{-\nu}}{\log 2} \max\left\{ \frac{D_{\rho}(\Psi) + 1 - \rho}{\theta}, D_1(\Psi) \right\}. \quad (177)$$

The analysis of this term is simpler than in the reverse Z-channel case. We again use Lemma 7, taking $\gamma_1 = \rho$, $\gamma_2 = 1$, $c = \theta$ and $d = 1 - \rho$. In this case, we find that (138) does not hold, since

$$1 - \rho \geq \theta(1-\rho) \geq \theta(1 - \rho + \rho\log\rho) = cD_1(\rho). \quad (178)$$

In other words, the maximum in (177) is always provided by the first term, and by choosing Ψ arbitrarily close to ρ, we deduce (also using $D_{\rho}(\rho) = 0$) that any rate above

$$\frac{(1-\theta)\nu e^{-\nu}}{\log 2} \frac{(1-\rho)}{\theta} \quad (179)$$

will ensure the error probability P_{e} does not converge to zero. Finally, using the fact that $\nu e^{-\nu} \leq 1/e$, we deduce that using any rate larger than

$$\frac{(1-\theta)(1-\rho)}{\theta e \log 2} \quad (180)$$

will ensure that P_{e} does not converge to zero.

F. Proof of Equation (129) in Lemma 5

It suffices to show that

$$\frac{\mathbb{P}\big[\mathrm{Binomial}\big(n, \frac{q_-}{1-q_+}\big) \geq d\big]}{\mathbb{P}\big[\mathrm{Binomial}\big(n, \frac{q_-}{1-2q_+}\big) \geq d\big]} \geq 1 + o(1). \quad (181)$$

To simplify the notation, we write P and Q for binomial PMFs (with n trials) with probabilities $p = \frac{q_-}{1-q_+}$ and $q = \frac{q_-}{1-2q_+}$. We are interested in lower bounding $\frac{\sum_{x\geq d} P(x)}{\sum_{x\geq d} Q(x)}$.

[10] For the first term, this is established by substituting (176) into the definition of $D_{(\cdot)}$ to obtain a ratio of the form $\frac{-\log\rho + o(\log\rho)}{-\log\rho + o(\log\rho)}$. For the second term, simply note that $\zeta e^{\nu} \to 1$ and $\alpha \to 0$.

To do this, we write $\sum_{x\geq d} Q(x) = \sum_{x\geq d} P(x) + \sum_{x\geq d}\big(Q(x) - P(x)\big)$ and seek to show that

$$\sum_{x\geq d}\big(Q(x) - P(x)\big) \ll \sum_{x\geq d} P(x), \quad (182)$$

in the sense that the left-hand side is upper bounded by a vanishing fraction of the right-hand side.

Towards establishing (182), note that

$$\sum_{x\geq d}\big(Q(x) - P(x)\big) = \sum_{x\geq d} P(x)\left(\frac{Q(x)}{P(x)} - 1\right) \quad (183)$$

$$= \sum_{x\geq d} P(x)\left(\left(\frac{q}{p}\right)^{x}\left(\frac{1-q}{1-p}\right)^{n-x} - 1\right). \quad (184)$$

Observe that under the choices $p = \frac{q_-}{1-q_+}$ and $q = \frac{q_-}{1-2q_+}$, we have

$$\frac{p}{q} = \frac{1-2q_+}{1-q_+} = 1 - \frac{q_+}{1-q_+} < 1. \quad (185)$$

We also have $\frac{1-p}{1-q} > 1$ (since $p < q$), so we can upper bound $\left(\frac{1-q}{1-p}\right)^{n-x}$ in (184) by one. Combining these findings gives

$$\sum_{x\geq d}\big(Q(x) - P(x)\big) \leq \sum_{x\geq d} P(x)\left(\left(\frac{q}{p}\right)^{x} - 1\right) \quad (186)$$

$$= \sum_{x\geq d} P(x)\left(\left(\frac{1}{1 - \frac{q_+}{1-q_+}}\right)^{x} - 1\right) \quad (187)$$

$$= \sum_{x\geq d} P(x)\big(e^{xq_+(1+o(1))} - 1\big), \quad (188)$$

since q_- and q_+ are both $o(1)$.

To complete the proof, recall that $d = O(\log p)$, whereas $q_+ = O\big(\frac{1}{k}\big)$ with $k \asymp p^{\theta}$ and $\theta > 0$. We fix $\xi < \theta$ and split the summation in (188) as follows:

$$\sum_{x\geq d} P(x)\big(e^{xq_+(1+o(1))} - 1\big) = \sum_{d\leq x\leq p^{\xi}} P(x)\big(e^{xq_+(1+o(1))} - 1\big) + \sum_{p^{\xi}\leq x\leq n} P(x)\big(e^{xq_+(1+o(1))} - 1\big). \quad (189)$$

The first summation behaves as $o\big(\sum_{x\geq d} P(x)\big)$, since the conditions $x \leq p^{\xi}$, $q_+ = O\big(\frac{1}{k}\big)$, $k \asymp p^{\theta}$, and $\xi \in (0,\theta)$ collectively imply $e^{xq_+(1+o(1))} = 1 + o(1)$. For the second summation, we recall that $P(x)$ is the PMF of a $\mathrm{Binomial}(n, q_-/(1-q_+))$ random variable, and make use of the concentration bound (19). Equating x with $nq_-/(1-q_+)(1+\epsilon)$ gives $\epsilon = \frac{x(1-q^+)}{nq_-} - 1$, which yields $\epsilon = \Omega\big(\frac{p^{\xi}}{\log p}\big)$ due to the fact that $nq_- \asymp \log p$ and $x \geq p^{\xi}$.

The concentration bound in (19) gives an exponent of $\frac{nq_-}{1-q_+}D_1(1+\epsilon)$, and we have $D_1(1+\epsilon) \asymp \epsilon\log\epsilon$ when $\epsilon \to \infty$; combining these observations yields an exponent on the order of $x\log\epsilon$, or equivalently, on the order of $x\log p$. Substituting $P(x) \le e^{-\Omega(x\log p)}$ into the second summation of (189), we readily obtain $\sum_{p^\xi \le x \le n} P(x)\big(e^{xq_+(1+o(1))}-1\big) \le e^{-\Omega(p^\xi \log p)}$. This is much faster decay than the $p^{-O(1)}$ decay of $\sum_{x\ge d} P(x)$ shown following (119).

Recalling that (189) is equal to (188), we deduce that (182) holds with $p = \frac{q_-}{1-q_+}$ and $q = \frac{q_-}{1-2q_+}$, as desired.

G. DD-Specific Converse Under RZ Noise

Consider the reverse Z-channel with parameter ρ. We claim that under i.i.d. Bernoulli testing with parameter $\nu > 0$, if the number of tests satisfies

$$n = \frac{k\log\frac{p}{k}}{\nu e^{-\nu}(1-\rho)}(1-\eta) \tag{190}$$

for some fixed $\eta \in (0,1)$, then the error probability of DD tends to one. We proceed by proving this claim.

First step. Recall that the first step removes all items that appear in any negative test. Given that there are $n_{\rm neg} = \big(ne^{-\nu}(1-\rho)\big)(1+o(1))$ negative tests (*cf.*, (49)), the probability of a given non-defective j being kept is

$$\psi := \left(1-\frac{\nu}{k}\right)^{n_{\rm neg}} = e^{-\left(\frac{n}{k}\nu e^{-\nu}(1-\rho)\right)(1+o(1))}. \tag{191}$$

Given $n_{\rm neg}$, the number of intruding non-defectives (i.e., non-defectives appearing in no negative tests) is distributed as $G \sim \mathrm{Binomial}(p-k,\psi)$, so if $\psi p \to \infty$ then we have by the binomial concentration bounds (19)–(20) that $G = p\psi(1+o(1))$ with probability approaching one. Using (190) and (191), we obtain

$$\psi = e^{-(1-\eta)\log\frac{p}{k}\cdot(1+o(1))} = \left(\frac{k}{p}\right)^{(1-\eta)(1+o(1))}, \tag{192}$$

and hence

$$p\psi = k^{(1-\eta)(1+o(1))}\cdot p^{\eta(1+o(1))}. \tag{193}$$

Since $k \asymp p^\theta$ with $\theta \in (0,1)$, we deduce that $p\psi \ge 2k^{1+\epsilon}$ for sufficiently large p and some $\epsilon > 0$ depending only on θ. Then, since $G = p\psi(1+o(1))$ with probability approaching one, we conclude that $G \ge k^{1+\epsilon}$ with probability approaching one.

Second step. We proceed similarly to the argument following Lemma 2, but now the key probability γ analyzed in (59) behaves very differently due to the fact that $g \ge k^{1+\epsilon}$. Specifically, we have

$$\gamma = \left(1-\frac{\nu}{k}\right)^g \tag{194}$$

$$= \left(\left(1-\frac{\nu}{k}\right)^{\frac{k}{\nu}}\right)^{\frac{\nu g}{k}} \tag{195}$$

$$= \Big(\big(e^{-1}\big)(1+o(1))\Big)^{\frac{\nu g}{k}} \tag{196}$$

$$\le e^{-\nu k^\epsilon(1+o(1))}, \tag{197}$$

where we have applied $g \ge k^{1+\epsilon}$. Hence, in the binomial distribution found in (62), the associated probability decreases by this factor accordingly:

$$(N'_{{\rm pos},j} \,|\, \widehat{\rm pd}, g, n_{\rm neg}) \sim \mathrm{Binomial}\Bigg(n(1-q_{\rm neg})(1+o(1)), \frac{\nu e^{-\nu\gamma}}{k(1-q_{\rm neg})}(1+o(1))\Bigg). \tag{198}$$

By (197), the mean of this distribution is at most $\frac{\nu n}{k}e^{-\nu k^\epsilon(1+o(1))}$, which vanishes under the choice of n in (190) due to the fact that $k \asymp p^\theta$ with $\theta > 0$. By Markov's inequality, it follows that the probability of the given defective item being the unique element from $\widehat{\rm pd}$ in *any* positive test vanishes as $p \to \infty$, and hence the probability that the DD algorithm successfully identifies a given defective item also vanishes.

H. High-ρ Low-θ Optimality Result for Reverse Z-Channel

In this appendix, we consider the alternative formulation of the achievability part of Theorem 1 given in Theorem 3. We let ξ be arbitrarily close to θ, and set $\beta = \frac{1+\rho}{2} \in (\rho,1)$. The conditions (44) and (45) both have a dependence on (k,p,θ) scaling as $O(k\log k) = O(\theta k\log p)$, which is dominated by the $k\log p$ term in (43) for sufficiently small θ. Hence, the condition (43) dominates for sufficiently small θ.

When (43) dominates, the required number of tests simplifies to

$$n = \left(\frac{e}{1-\rho}\cdot k\log p\right)(1+o(1)), \tag{199}$$

which yields a rate of $\frac{1-\rho}{e\log 2}$ bits per test. By performing a Taylor expansion of (11) at $\rho = 1$, we find that the reverse Z-channel capacity also behaves as $C = \frac{1-\rho}{e\log 2} + O\big((1-\rho)^2\big)$ as $\rho \to 1$, so in fact we have asymptotic optimality in this high-noise regime.

Stated more precisely, the rate of Theorem 3 is asymptotically optimal when the order of limits is first $n \to \infty$, then $\theta \to 0$, and finally $\rho \to 1$.

I. Results for the General Binary Noise Model

Achievability. Under the general binary noise model (which includes symmetric noise as a special case), we consider the following noisy version of the DD algorithm.

Noisy DD algorithm for general binary noise:

1. For each $j \in [p]$, let $N_{\mathrm{neg},j}$ be the number of negative tests in which item j is included. In the first step, we fix a constant $\alpha \in (\rho_{10}, 1-\rho_{01})$ and construct the following set of items that are believed to be non-defective:
$$\widehat{\mathrm{ND}} = \left\{ j \in [p] \,:\, N_{\mathrm{neg},j} \geq \frac{\alpha n \nu}{k} \right\}. \tag{200}$$
The remaining items, $\widehat{\mathrm{PD}} = [p] \setminus \widehat{\mathrm{ND}}$, are believed to be "possibly defective".
2. For each $j \in \widehat{\mathrm{PD}}$, let $N'_{\mathrm{pos},j}$ be the number of positive tests that include item j and no other item from $\widehat{\mathrm{PD}}$. In the second step, we fix a constant $\beta \in (\rho_{01}, 1-\rho_{10})$, and estimate the defective set as follows:
$$\widehat{S} = \left\{ j \in \widehat{\mathrm{PD}} \,:\, N'_{\mathrm{pos},j} \geq \frac{\beta n \nu e^{-\nu}}{k} \right\}. \tag{201}$$

Theorem 5. (General binary noise achievability) *Consider the general binary noisy group testing setup with crossover probabilities ρ_{01} and ρ_{10} (cf., (7)), number of defectives $k \asymp p^\theta$ (where $\theta \in (0,1)$), and i.i.d. Bernoulli testing with parameter $\nu > 0$. For any $\alpha \in (\rho_{10}, 1-\rho_{01})$, $\beta \in (\rho_{01}, 1-\rho_{10})$, and $\xi \in (0,\theta)$, we have $P_{\mathrm{e}} \to 0$ if*
$$n \geq \max\left\{ n_1^{(\mathrm{D})}, n_1^{(\mathrm{ND})}, n_2^{(\mathrm{D})}, n_2^{(\mathrm{ND})} \right\}(1+\eta) \tag{202}$$
for arbitrarily small $\eta > 0$, where defining $w = (1-\rho_{01})e^{-\nu} + \rho_{10}(1-e^{-\nu})$, we have
$$n_1^{(\mathrm{D})} = \frac{1}{\nu \rho_{10} D_1(\alpha/\rho_{10})} \cdot k \log k, \tag{203}$$
$$n_1^{(\mathrm{ND})} = \frac{1-\xi}{\nu w D_1(\alpha/w)} \cdot k \log p, \tag{204}$$
$$n_2^{(\mathrm{D})} = \frac{1}{\nu e^{-\nu}(1-\rho_{10}) D_1(\beta/(1-\rho_{10}))} \cdot k \log k, \tag{205}$$
$$n_2^{(\mathrm{ND})} = \frac{\xi}{\nu e^{-\nu} \rho_{01} D_1(\beta/\rho_{01})} \cdot k \log p. \tag{206}$$

Proof. The first step of the algorithm is analyzed in the same way as the proof of Theorem 4, and the second step is analyzed in the same way as the proofs of Theorems 3 and 4. For the former, $\widehat{\mathrm{PD}}$ contains all k defective items and $o(k)$ non-defective items under the conditions (203)–(204), and for the latter, $\widehat{S}$ contains all of the defective items and no non-defective items under the conditions (205)–(206). The details are omitted to avoid repetition. □

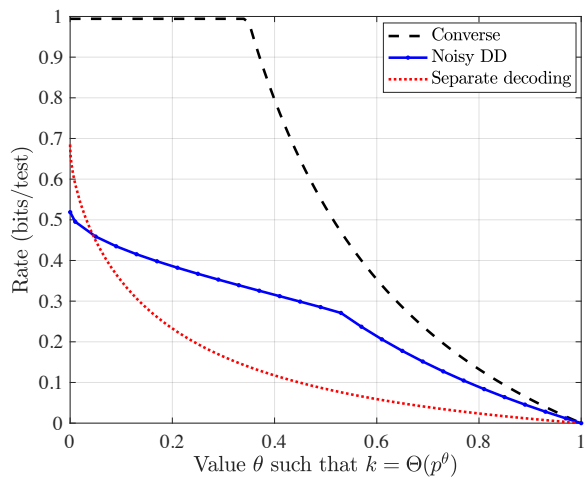

Figure 7: Achievable and converse rates under the symmetric noise model and i.i.d. Bernoulli testing with noise level $\rho = 0.001$.

As before, we can rephrase Theorem 5 to give a statement in terms of rates. That is, we can rewrite (203), (204), (205) and (206) using (15) to show that for given α, β, ξ, and ν an achievable rate is:
$$\frac{1-\theta}{\log 2} \min\left\{ \frac{\nu}{\theta} D_{\rho_{01}}(\alpha), \frac{\nu}{1-\xi} D_w(\alpha), \frac{\nu e^{-\nu}}{\theta} D_{1-\rho_{10}}(\beta), \frac{\nu e^{-\nu}}{\xi} D_{\rho_{01}}(\beta) \right\}. \tag{207}$$

We claim that in the limit as $\rho_{01} \to 0$ and $\rho_{10} \to 0$ for fixed $\theta \in (0,1)$, we recover the noiseless DD guarantee given in [12]. To see this, we let both α and β equal an arbitrarily small constant $c > 0$. We observe that $D_1(\alpha/\rho)$ and $D_1(\beta/\rho)$ both scale as $\frac{1}{\rho}\log\frac{1}{\rho}$, which implies that $n_1^{(\mathrm{D})}$ and $n_2^{(\mathrm{ND})}$ behave as $o(k \log k)$, and their contributions are insignificant. On the other hand, α/w and $\beta/(1-\rho_{10})$ can be made arbitrarily close to zero by suitable choice of c, which means (using $D_1(0) = 1$ and $w|_{\rho_{01}=\rho_{10}=0} = e^{-\nu}$) that $n_1^{(\mathrm{ND})}$ and $n_2^{(\mathrm{D})}$ can be made arbitrarily close to $(1-\xi)\frac{k \log p}{\nu e^{-\nu}}$ and $\frac{k \log k}{\nu e^{-\nu}}$, respectively. Since ξ can be chosen arbitrarily close to θ and $(1-\theta)k \log p = \big(k \log \frac{p}{k}\big)(1+o(1))$, the final condition on n is
$$n \geq \max\left\{ \frac{k \log \frac{p}{k}}{\nu e^{-\nu}}, \frac{k \log k}{\nu e^{-\nu}} \right\}(1+\eta) \tag{208}$$
for arbitrarily small $\eta > 0$. This bound is minimized by the choice $\nu = 1$, which recovers the bound in [12].

Converse. We can construct a general binary channel with noise levels $\rho_{01} + \rho_{10} \leq 1$ by the composition of a reverse Z-channel followed by a Z-channel. To be precise, analyzing the relevant conditional probabilities, using a reverse Z-channel with noise level $\rho_{01}/(1-\rho_{10}) \leq 1$

followed by a Z-channel with noise level ρ_{10} gives an general binary channel with noise levels ρ_{01} and ρ_{10}.

Hence, a standard genie argument shows that any converse that applies to the reverse Z-channel must also apply to the general binary channel. In particular, evaluating the converse bound $\overline{R}^{(\mathrm{RZ})}\left(\theta, \rho_{01}/(1-\rho_{10})\right)$ of Theorem 1 gives a converse that applies to the general binary channel.

We can use a similar argument based on taking a Z-channel with noise level $\rho_{10}/(1-\rho_{01}) \leq 1$ followed by a reverse Z-channel with noise level ρ_{01}. The same genie argument shows that we can bound the rate of the general binary channel by $\overline{R}^{(\mathrm{Z})}\left(\theta, \rho_{10}/(1-\rho_{01})\right)$. Putting these bounds together, we deduce the following.

Corollary 1. *For the general binary noisy group testing problem with noise levels $\rho_{01}+\rho_{10} \leq 1$, in the regime $k \asymp p^{\theta}$ with $\theta \in (0,1)$, no algorithm can achieve $P_e \to 0$ under Bernoulli testing with a rate higher than*

$$\begin{aligned}&\overline{R}^{(\mathrm{gen})}(\theta, \rho_{01}, \rho_{10}) \\ &= \min\left\{\overline{R}^{(\mathrm{RZ})}\left(\theta, \frac{\rho_{01}}{1-\rho_{10}}\right), \overline{R}^{(\mathrm{Z})}\left(\theta, \frac{\rho_{10}}{1-\rho_{01}}\right)\right\}.\end{aligned} \tag{209}$$

In the degenerate case of $\rho_{01}+\rho_{10}=1$ the initial reverse Z-channel has a noise level of 1, meaning that all inputs are deterministically mapped to 1 by this step. In this case, as expected, the converse bound (209) is given by $\overline{R}^{(\mathrm{RZ})}(\theta,1)=0$ (where this value follows by (28)).

For comparison purposes, we discuss the symmetric case $\rho_{01}=\rho_{10}=\rho$. In this case, while Corollary 1 does not exactly match Theorem 5 for any $\theta \in (0,1)$ and $\rho \in (0,1)$, the two become increasingly close for high θ and low ρ; see Figure 7 for an example with $\rho = 0.001$. In this figure, we also observe a strict improvement over the best previously known rate attained by separate decoding of items [19] unless θ is very small.

The difficulty in establishing a tight bound when both ρ_{01} and ρ_{10} are positive appears to stem from the requirement of handling all four possible error types (false positive vs. false negative, and first stage vs. second stage); in contrast, for the Z and RZ models, only a strict subset of these is relevant. The fact that these error events do not occur independently of one another poses a significant challenge for a tight analysis. Similarly, in the converse proof, constructing a set $S \setminus \{i\} \cup \{j\}$ with a higher likelihood than S is complicated by the fact that the likelihood depends on all four combinations of $P_{Y|U}(y|u)$ with $u, y \in \{0,1\}$. Further closing the remaining gaps remains an interesting direction for future research.

Acknowledgment

J. Scarlett was supported by an NUS Startup Grant.

References

[1] M. Aldridge, O. Johnson, and J. Scarlett, "Group testing: An information theory perspective," *Found. Trend. Comms. Inf. Theory*, vol. 15, no. 3–4, pp. 196–392, 2019.

[2] R. Dorfman, "The detection of defective members of large populations," *Ann. Math. Stats.*, vol. 14, no. 4, pp. 436–440, 1943.

[3] A. Fernández Anta, M. A. Mosteiro, and J. Ramón Muñoz, "Unbounded contention resolution in multiple-access channels," in *Distributed Computing*. Springer Berlin Heidelberg, 2011, vol. 6950, pp. 225–236.

[4] Y. Erlich, A. Gilbert, H. Ngo, A. Rudra, N. Thierry-Mieg, M. Wootters, D. Zielinski, and O. Zuk, "Biological screens from linear codes: Theory and tools," 2015, https://www.biorxiv.org/content/10.1101/035352v1.article-info.

[5] M. T. Goodrich, M. J. Atallah, and R. Tamassia, "Indexing information for data forensics," in *Conf. Appl. Crypt. Net. Security*. Springer, 2005, pp. 206–221.

[6] R. Clifford, K. Efremenko, E. Porat, and A. Rothschild, "Pattern matching with don't cares and few errors," *J. Comp. Sys. Sci.*, vol. 76, no. 2, pp. 115–124, 2010.

[7] G. Cormode and S. Muthukrishnan, "What's hot and what's not: Tracking most frequent items dynamically," *ACM Trans. Database Sys.*, vol. 30, no. 1, pp. 249–278, March 2005.

[8] A. Gilbert, M. Iwen, and M. Strauss, "Group testing and sparse signal recovery," in *Asilomar Conf. Sig., Sys. and Comp.*, Oct. 2008, pp. 1059–1063.

[9] A. C. Gilbert, M. J. Strauss, J. A. Tropp, and R. Vershynin, "One sketch for all: Fast algorithms for compressed sensing," in *Proc. ACM-SIAM Symp. Disc. Alg. (SODA)*, New York, 2007, pp. 237–246.

[10] G. Atia and V. Saligrama, "Boolean compressed sensing and noisy group testing," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1880–1901, March 2012.

[11] L. Baldassini, O. Johnson, and M. Aldridge, "The capacity of adaptive group testing," in *IEEE Int. Symp. Inf. Theory*, July 2013, pp. 2676–2680.

[12] M. Aldridge, L. Baldassini, and O. Johnson, "Group testing algorithms: Bounds and simulations," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3671–3687, June 2014.

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

[14] M. Aldridge, "The capacity of Bernoulli nonadaptive group testing," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7142–7148, 2017.

[15] O. Johnson, M. Aldridge, and J. Scarlett, "Performance of group testing algorithms with near-constant -item," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 707–723, Feb. 2019.

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

[17] C. L. Chan, S. Jaggi, V. Saligrama, and S. Agnihotri, "Non-adaptive group testing: Explicit bounds and novel algorithms," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 3019–3035, May 2014.

[18] T. Laarhoven, "Asymptotics of fingerprinting and group testing: Tight bounds from channel capacities," *IEEE Trans. Inf. Forens. Sec.*, vol. 10, no. 9, pp. 1967–1980, 2015.

[19] J. Scarlett and V. Cevher, "Near-optimal noisy group testing via separate decoding of items," *IEEE Trans. Sel. Topics Sig. Proc.*, vol. 2, no. 4, pp. 625–638, 2018.

[20] J. Scarlett, "Noisy adaptive group testing: Bounds and algorithms," *IEEE Trans. Inf. Theory*, vol. 65, no. 6, pp. 3646–3661, June 2019.

[21] F. K. Hwang, "A method for detecting all defective members in a population by group testing," *J. Amer. Stat. Assoc.*, vol. 67, no. 339, pp. 605–608, 1972.

[22] M. Aldridge, L. Baldassini, and K. Gunderson, "Almost separable matrices," *J. Comb. Opt.*, pp. 1–22, 2015.

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

[24] L. G. Tallini, S. Al-Bassam, and B. Bose, "On the capacity and codes for the z-channel," in *IEEE Int. Symp. Inf. Theory*, 2002.

[25] M. B. Malyutov and P. S. Mateev, "Screening designs for non-symmetric response function," *Mat. Zametki*, vol. 29, pp. 109–127, 1980.

[26] M. Malyutov, "The separating property of random matrices," *Math. Notes Acad. Sci. USSR*, vol. 23, no. 1, pp. 84–91, 1978.

[27] J. Scarlett and V. Cevher, "Limits on support recovery with probabilistic models: An information-theoretic framework," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 593–620, 2017.

[28] ——, "Converse bounds for noisy group testing with arbitrary measurement matrices," in *IEEE Int. Symp. Inf. Theory*, Barcelona, 2016.

[29] D. Sejdinovic and O. Johnson, "Note on noisy group testing: Asymptotic bounds and belief propagation reconstruction," in *Allerton Conf. Comm., Control and Comp.*, 2010.

[30] D. Malioutov and M. Malyutov, "Boolean compressed sensing: LP relaxation for group testing," in *IEEE Int. Conf. Acoust. Sp. Sig. Proc. (ICASSP)*, March 2012, pp. 3305–3308.

[31] S. Cai, M. Jahangoshahi, M. Bakshi, and S. Jaggi, "Efficient algorithms for noisy group testing," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 2113–2136, 2017.

[32] K. Lee, R. Pedarsani, and K. Ramchandran, "SAFFRON: A fast, efficient, and robust framework for group testing based on sparse-graph codes," 2015, http://arxiv.org/abs/1508.04485.

[33] H. A. Inan, P. Kairouz, M. Wootters, and A. Özgür, "On the optimality of the Kautz-Singleton construction in probabilistic group testing," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5592–5603, Sept. 2019.

[34] S. Bondorf, B. Chen, J. Scarlett, H. Yu, and Y. Zhao, "Sublinear-time non-adaptive group testing with $O(k \log n)$ tests via bit-mixing coding," 2019, https://arxiv.org/abs/1904.10102.

[35] M. Cheraghchi, "Noise-resilient group testing: Limitations and constructions," *Disc. App. Math.*, vol. 161, no. 1, pp. 81–95, 2013.

[36] H. Q. Ngo, E. Porat, and A. Rudra, "Efficiently decodable error-correcting list disjunct matrices and applications," in *Int. Colloq. Automata, Lang., and Prog.*, 2011.

[37] P. Indyk, H. Q. Ngo, and A. Rudra, "Efficiently decodable non-adaptive group testing," in *ACM-SIAM Symp. Disc. Alg. (SODA)*, 2010.

[38] A. J. Macula, "Error-correcting nonadaptive group testing with de-disjunct matrices," *Disc. App. Math.*, vol. 80, no. 2-3, pp. 217–222, 1997.

[39] T. Madej, "An application of group testing to the file comparison problem," in *Distributed Computing Systems, 1989., 9th International Conference on*. IEEE, 1989, pp. 237–243.

[40] J. K. Wolf, "Born again group testing: Multiaccess communications," *IEEE Trans. Inf. Theory*, vol. 31, no. 2, pp. 185–191, 1985.

[41] T. Berger and V. I. Levenshtein, "Asymptotic efficiency of two-stage disjunctive testing," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1741–1749, 2002.

[42] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Adv. Comp. Math.*, vol. 5, no. 1, pp. 329–359, 1996.

[43] R. Motwani and P. Raghavan, *Randomized Algorithms*. Chapman & Hall/CRC, 2010.

[44] R. B. Ash, *Information Theory*. Dover Publications Inc., New York, 1990.

[45] K. Joag-Dev and F. Proschan, "Negative association of random variables with applications," *Ann. Stats.*, vol. 11, pp. 286–295, 1983.

Jonathan Scarlett (S'14 – M'15) received the B.Eng. degree in electrical engineering and the B.Sci. degree in computer science from the University of Melbourne, Australia. From October 2011 to August 2014, he was a Ph.D. student in the Signal Processing and Communications Group at the University of Cambridge, United Kingdom. From September 2014 to September 2017, he was post-doctoral researcher with the Laboratory for Information and Inference Systems at the École Polytechnique Fédérale de Lausanne, Switzerland. Since January 2018, he has been an assistant professor in the Department of Computer Science and Department of Mathematics, National University of Singapore. His research interests are in the areas of information theory, machine learning, signal processing, and high-dimensional statistics. He received the Singapore National Research Foundation (NRF) fellowship, and the NUS Early Career Research Award.

Oliver Johnson received the B.A. degree in 1995, Part III Mathematics in 1996, and the Ph.D. degree in 2000, all from the University of Cambridge, Cambridge, U.K. He was Clayton Research Fellow at Christ's College and Max Newman Research Fellow at Cambridge University until 2006, during which time he published the book Information Theory and the Central Limit Theorem (Singapore: World Scientific, 2004). Since 2006, he has been at University of Bristol, Bristol, U.K, where he is Professor of Information Theory. Within Bristol Mathematics he is Director of the Institute for Statistical Science and Programme Director for the MSc in the Mathematics of Cybersecurity.