Asymptotic Efficiency of Two-Stage Disjunctive Testing

Toby Berger, *Fellow, IEEE,* and Vladimir I. Levenshtein, *Member, IEEE*

***Abstract*—We adapt methods originally developed in information and coding theory to solve some testing problems. The efficiency of two-stage pool testing of n items is characterized by the minimum expected number $E(n, p)$ of tests for the Bernoulli p-scheme, where the minimum is taken over a matrix that specifies the tests that constitute the first stage. An information-theoretic bound implies that the natural desire to achieve $E(n, p) = o(n)$ as $n \to \infty$ can be satisfied only if $p(n) \to 0$. Using random selection and linear programming, we bound some parameters of binary matrices, thereby determining up to positive constants how the asymptotic behavior of $E(n, p)$ as $n \to \infty$ depends on the manner in which $p(n) \to 0$. In particular, it is shown that for $p(n) = n^{-\beta+o(1)}$, where $0 < \beta < 1$, the asymptotic efficiency of two-stage procedures cannot be improved upon by generalizing to the class of all multistage adaptive testing algorithms.**

***Index Terms*—Cover-free codes, disjunctive testing, linear programming, pool testing, random selection, reconstruction algorithms, screening.**

I. Introduction

We study the theory and design of efficient combinatorial and probabilistic pool testing procedures. The object of pool testing is to identify an *a priori* unknown subset of $N_n = \{1, \ldots, n\}$ called the set of *active* items using as few queries as possible. Each query informs the tester about whether or not a certain subset of N_n called a *pool* has a nonempty intersection with the set of active items. A negative answer to this question gives information that all items belonging to the pool are inactive. This approach has been used in many applications beginning with an efficient blood testing problem in [7]. Other applications include (following [13] and [10]) quality control in product testing [22], searching files in storage systems [12], efficient accessing of computer memories [12], sequential screening of experimental variables [16], efficient contention resolution algorithms for multiple-access communications, [3], [23], [17], and screening of clone libraries [2], [4]. The books and review papers [6], [1], [8], and [13] also are concerned with this topic. In this investigation, we use traditional information-theoretic methods and emphasize two-stage testing because of its importance in modern biological applications such as monoclonal antibody generation and cDNA library screening.

We consider only "gold standard" tests characterized by zero false positives (i.e., unit sensitivity) and zero false negatives (i.e., unit specificity). In practice, of course, false positives and false negatives occur. In cDNA library screening, however, polymerase chain reaction (PCR) amplification techniques provide tests whose reliability closely approximates that of gold standard tests. In any event, determining the optimum efficiency attainable with gold standard tests provides an absolute standard with which to compare and a goal toward which to strive. For an approach to analysis of testing in the face of false positives and negatives, see [13], [18].

There are many families of algorithms designed to ascertain the value $(x_1, \ldots, x_n)$ that has been assumed by an *a priori* unknown vector X via application to this vector a succession of permissible operations (tests). Among these, pool testing algorithms are those algorithms in which the only permissible operations are pool tests as defined above. In general, the structure of the next test depends on the results of previous tests, in which case we say the algorithm is *adaptive*. Efficient reconstruction of X is connected with minimization of the number of tests. Given a probability distribution governing selection of X, the expected number of tests required to ascertain the value $(x_1, \ldots, x_n)$ that X assumes depends, of course, on which test types are permissible. However, there exists a general information-theoretic bound which depends only on the cardinality, call it q, of the range of the tests and the probability distribution of X. This bound is a direct consequence of Shannon's theorem on q-ary prefix coding. First, we formulate this bound for general reconstruction algorithms and then consider in detail algorithms for reconstruction of binary vectors based on disjunctive tests. In a recent paper [14], one of the authors investigates another problem of reconstructing an unknown vector using the minimum number of boundedly distorted versions thereof.

In Section II, we present definitions and notations needed to set the problem in more detail and then describe how the remainder of the paper is organized and the main results that are obtained.

II. Reconstruction Algorithms

We denote by H_r^n the set all vectors $X = (x_1, \ldots, x_n)$ over the alphabet $H_r = \{0, 1, \ldots, r-1\}$, $r \geq 2$. We also write $X = x_1 \cdots x_n$, considering X as a word of length n over H_r, and put $H_r^* = \bigcup_{n=0}^{\infty} H_r^n$; here, H_r^0 is a singleton containing the

Manuscript received August 8, 2001; revised February 14, 2002. The work of T. Berger was supported by the NSF under Grants ECS-9632266 and CCR-9980616. the work of V. I. Levenshtein was supported in part by the Russian Foundation for Basic Research under Grant 01-01-00035.

T. Berger is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: berger@ece.cornell.edu).

V. I. Levenshtein is with the Keldysh Institute for Applied Mathematics, RAS, Miusskaya Sq. 4, 125047 Moscow, Russia (e-mail: leven@spp.keldysh.ru).

Communicated by G. Battail, Associate Editor At Large.

Publisher Item Identifier S 0018-9448(02)05153-2.

empty word. We assume that $X \in H_r^n$ is selected according to a probability distribution P. This means that we, in fact, consider the reconstruction problem for the set

$$H_r^n(P) = \{X \in H_r^n \colon P(X) > 0\}. \tag{1}$$

This formulation allows us to consider not only probabilistic problems, but also combinatorial problems in which it is known only that X belongs to some specified subset $H \subseteq H_r^n$. We call a probability distribution *symmetric* if $P(X)$ does not change for any permutation of the components of X. The Bernoulli distribution on H_2^n, which assigns to any $X \in H_2^n$ composed of i ones and $n-i$ zeros the probability $P(X) = p^i(1-p)^{n-i}$ where $0 < p < 1$, $i = 0, 1, \ldots, n$, is an example of symmetric distribution P for which $H_2^n(P) = H_2^n$. We call this distribution the *Bernoulli p-scheme*.

Any *reconstruction algorithm for $H_r^n(P)$ over H_q* based on successive application of tests can be described by a partial function $F(X, S)$ in two variables

$$F \colon H_r^n \times H_q^* \to H_q \tag{2}$$

satisfying a property that follows. The word $S = s_1 \cdots s_l \in H_q^l$ is called a (current) *syndrome* for X (a connection with syndromes for linear codes will be explained later) if

$$s_i = F(X, s_1 \cdots s_{i-1}), \qquad \text{for and } i = 1, \ldots, l. \tag{3}$$

(Here $s_1 \cdots s_{i-1}$ is the empty word for $i = 1$.) The syndrome $S = s_1 \cdots s_l$ describes results of the first l tests applied to X. A syndrome S for $X \in H_r^n(P)$ is called *complete* if it is not a syndrome for any $Y \in H_r^n(P)$, $Y \neq X$. We assume that $F(X, S)$ has the following property: for any $X \in H_r^n(P)$ there exists a complete syndrome. This means that utilizing algorithm F, which consists of successive application of $F(X, S)$, enables one to reconstruct any $X \in H_r^n(P)$. (The value $F(X, S)$ may not be defined if a proper prefix of S is a complete syndrome for X.) The expected number $\bar{l}(H_r^n(P), q)$ of tests in this reconstruction algorithm is defined by

$$\bar{l}(H_r^n(P), q) = \sum_{X \in H_r^n(P)} P(X) l(X) \tag{4}$$

where $l(X)$ is the minimum length of a complete syndrome for X. Note that minimum-length complete syndromes for all $X \in H_r^n(P)$ form a prefix code over the alphabet H_q. Therefore, (4) coincides with the expected length of the prefix code comprising the minimum-length complete syndromes, and we have the following consequence of Shannon's theorem on prefix coding [21].

The Information-Theoretic Bound: Given a probability distribution P on H_r^n, for any reconstruction algorithm for $H_r^n(P)$ over H_q

$$\bar{l}(H_r^n(P), q) \geq -\sum_{X \in H_r^n(P)} P(X) \log_q P(X). \tag{5}$$

Specifically, in the case of the Bernoulli p-scheme

$$\bar{l}(H_2^n, 2) \geq n\left(-p\log_2 p - (1-p)\log_2(1-p)\right). \tag{6}$$

The value of $\bar{l}(H_r^n(P), q)$ depends critically on what restrictions are imposed on the function $F(X, S)$ and/or on the form of permissible tests. An important case considered in the paper is characterized by a function

$$G \colon H_r^n \times H_q^* \to H_r^n \tag{7}$$

such that the result of the test $F(X, S)$ is defined as an "inner product" of vectors

$$s_i = F(X, s_1 \cdots s_{i-1}) = \langle X, Z \rangle = \sum_{i=1}^{n} x_i \cdot z_i, \tag{8}$$

where

$$\begin{aligned} X &= (x_1, \ldots, x_n) \\ Z &= (z_1, \ldots, z_n) = G(X, s_1 \cdots s_{i-1}) \end{aligned}$$

and the calculations are performed with respect to some choice of summation and multiplication operations defined on H_r^n. In particular, in the case $r = 2$, the operations could be either real-field summation and multiplication, or summation and multiplication in the field GF(2), or the logical operations disjunction $\vee$ and conjunction $\wedge$. These choices give rise, respectively, to the known problems of finding counterfeit coins on an accurate scale [9], finding additive noises for linear codes [19], and finding active items using pool testing [12].

In this paper, we consider tests (8) for $r = q = 2$ and the third alternative

$$\langle X, Z \rangle = \bigvee_{i=1}^{n} x_i \wedge z_i.$$

For any $X = (x_1, \ldots, x_n) \in H_2^n$, we shall also denote by X the subset of $N_n = \{1, \ldots, n\}$ consisting of all i such that $x_i = 1$. Elements of the sets X and $N_n \setminus X$, respectively, are called the *active* and the *inactive* items of the vector X. Whether we want X to mean a vector or to mean a subset comprised of this vector's active items will be clear from the context. In particular, $\langle X, Z \rangle = 1$ if and only if the subset (pool) Z has nonempty intersection with the set X.

The simplest testing procedure, a *one-stage* algorithm, is defined by an $m \times n$ binary matrix $A = (a_{ij})$ with rows $A_i = (a_{i1}, \ldots, a_{im})$. The one-stage procedure consists of calculating for any $X \in H_2^n$ the syndrome $S = (s_1, \ldots, s_m) \in H_2^m$, where $s_i = \langle X, A_i \rangle$, $i = 1, \ldots, m$. We shall write it as $S = XA^T$, where T denotes transposition and the logical operations $\vee$ and $\wedge$ are used in the matrix product. (Formally, this algorithm is a special case of (8) wherein $G(X, S)$ depends only on the length of the word S.) For any vector $S \in H_2^m$, denote by $Q(A, S)$ the (possibly empty) set of all vectors $X \in H_2^n$, such that $S = XA^T$. If $X \in Q(A, S)$, then S is a syndrome for X which is complete if $|Q(A, S)| = 1$. For $j = 1, 2, \ldots, n$, let B_j denote both the jth column $(a_{1j}, \ldots, a_{mj}) \in H_2^m$ of A and also the set $\{i \colon a_{ij} = 1\} \subseteq N_m$ of its active items; this notational convention is wholly analogous to that which we have earlier introduced of letting $X \in H_2^n$ denote both the unknown vector and the set of so-called active items which index whichever components of X equal 1 as opposed to 0. Then this syndrome can be represented as

$$S = \bigcup_{j \in X} B_j. \tag{9}$$

Denote by $H_2^{n,\,t}$ the set of all $X \in H_2^n$ with exactly t active items and by $H_2^{n,\,\le t}$ the subset of all $X \in H_2^n$ with t or fewer active items. A matrix A (and its columns) is referred to as a *disjunctive t-code* if for any $S \in H_2^m$

$$\left| Q(A,\, S) \cap H_2^{n,\,\le t} \right| \le 1.$$

In other words, the syndromes of all $X \in H_2^{n,\,\le t}$ are distinct. Hence, the one-stage algorithm defined by this A enables one to reconstruct any unknown vector in $H_2^{n,\,\le t}$; note that this implies that it solves the combinatorial testing problem in which all vectors not in $H_2^{n,\,\le t}$ have probability 0. This is analogous to the fact that, if A is a check matrix of a linear t-error-correcting binary code (i.e., any $2t$ columns of A are linearly independent), then one can recover any error vector X from knowledge of its syndrome calculated using operations $\mathrm{mod}\, 2$ (as opposed to $\vee$ and $\wedge$ used in the present paper) provided the number of errors does not exceed t; see [19].

A matrix A (and its columns) is referred to as a *t-cover-free code* if for any X with t active items and any inactive item i $(i \in N_n \setminus X)$

$$B_i \not\subseteq \bigcup_{j \in X} B_j.$$

Disjunctive and cover-free codes were introduced in [12] where it was also shown that

$$t(A) \le t^-(A) \le t(A) + 1 \tag{10}$$

where $t^-(A)$ is the maximum number t such that a given matrix A without zero columns is a disjunctive t-code and $t(A)$ is the maximum t such that A is a t-cover-free code. There exists a trivial one-stage reconstruction algorithm for H_2^n which consists of an individual test for each of the n items (the matrix A is then the unit matrix or a permutation of it) and this number of tests cannot be decreased in the class of one-stage algorithms (see, for example, [5]).

In order to define two-stage reconstruction algorithms and describe their capabilities, we introduce additional terminology and notation. Fix a binary matrix A of size $m \times n$ with rows A_i and columns B_j, and a vector $S \in H_2^m$ such that $Q(A,\, S)$ is not empty. Note that if $X \in Q(A,\, S)$, then $S = XA^T$ and we have (9). Set

$$X(S) = \{j\colon j \in N_n,\, B_j \subseteq S\} \tag{11}$$

and recall our convention that $X(S)$ then also denotes the binary n-vector whose unit entries are in the positions $j \in X(S)$. Then we have $X(S) \in Q(A,\, S)$ and $X \subseteq X(S)$ for any $X \in Q(A,\, S)$.

What information about vectors in the set $Q(A,\, S)$ can we derive from their common syndrome S? An item $i \in N_n$ is called *negative* if $B_i \not\subseteq S$. By this definition, all negative items must be inactive in all vectors that belong to $Q(A,\, S)$. An item $i \in N_n$ is called *positive* if $B_i \subseteq S$ but

$$\bigcup_{j \in X(S) \setminus \{i\}} B_j \ne S.$$

By this definition, all positive items must be active in all vectors of $Q(A,\, S)$. The remaining items of N_n are called *unresolved*. This terminology is appropriate because, for any unresolved item $i \in N_n$, both $X(S)$ and $X(S) \setminus \{i\}$ belong to $Q(A,\, S)$. For each X, denote by $u(A,\, X)$ the number of unresolved items for the syndrome $S = XA^T$ for X, and let $\overline{X}$ be a compressed notation for the set $X(S) \subset N_n$ defined by (11). Since all items of $\overline{X} \setminus X$ are unresolved and $\overline{X}$ comprises all the positive and unresolved items, we have

$$|\overline{X} \setminus X| \le u(A,\, X) \le |\overline{X}|. \tag{12}$$

In Example 1, for the matrix A (whose columns are the Steiner triples) and syndrome $S = (111100111)$, we have $t(A) = 2$, $t^-(A) = 3$, $X(S) = \{1,\, 2,\, 5,\, 8,\, 12\}$, and all these items are unresolved; the remaining items $\{3,\, 4,\, 6,\, 7,\, 9,\, 10,\, 11\}$ are negative, and $Q(A,\, S)$ consists of 10 vectors whose sets of active items are subsets of $X(S)$ of which four are of cardinality 3, five of cardinality 4, and one of cardinality 5. Also, for any $X \in Q(A,\, S)$, $\overline{X} = X(S)$, $|\overline{X}| = 5$, and $|\overline{X} \setminus X| = 5 - |X| \in \{0,\, 1,\, 2\}$.

Example 1:

X	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	S^T
	1	1	1	1	0	0	0	0	0	0	0	0	1
	1	0	0	0	1	1	1	0	0	0	0	0	1
	1	0	0	0	0	0	0	1	1	1	0	0	1
	0	1	0	0	1	0	0	1	0	0	1	0	1
A	0	0	1	0	0	1	0	0	1	0	1	0	0
	0	0	0	1	0	0	1	0	0	1	1	0	0
	0	1	0	0	0	0	1	0	1	0	0	1	1
	0	0	0	1	0	1	0	1	0	0	0	1	1
	0	0	1	0	1	0	0	0	0	1	0	1	1

A *two-stage* reconstruction algorithm for H_2^n consists of: Stage 1—applying to the unknown $X \in H_2^n$ the m tests given by the rows of a fixed $m \times n$ matrix A, and Stage 2—resolving each of the $u(A,\, X)$ items left unresolved after Stage 1 by testing it individually. In many applications, it is possible to conduct all the tests of either stage simultaneously, which is what motivates the choice of the term "stage." The expected number $\tilde{l}(A,\, P)$ of tests in a two-stage reconstruction algorithm for H_2^n is

$$\tilde{l}(A,\, P) = m + \tilde{u}(A,\, P) \tag{13}$$

where

$$\tilde{u}(A,\, P) = \sum_{X \in H_2^n} P(X) u(A,\, X). \tag{14}$$

We write p instead of P in notations $\tilde{l}(A,\, P)$ and $\tilde{u}(A,\, P)$ in the case of the Bernoulli p-scheme.

The main aim of this paper is to investigate the minimum expected number of tests, namely,

$$E(n, p) = \min \tilde{l}(A, p) \tag{15}$$

where the minimum is taken over all binary matrices A with n columns. In [5], we noted that, for any matrix A and an unknown $X \in H_2^n$, one can determine from the syndrome $S = XA^T$ whether $|X| \le t(A)$ or $|X| > t(A)$ and reconstruct X in the first case. This gives rise to highly efficient two-stage testing which uses for its first stage a t-cover-free code with t slightly larger than the expected number np of active items. However, we shall verify that there exist more efficient two-stage testing algorithms based on proper selection of matrices A.

Inequality (6) shows that, unless $p(n) \to 0$ as $n \to \infty$, two-stage testing cannot affect an asymptotic improvement over one-stage testing's required n tests in the sense of achieving $E(n, p) = o(n)$. Our main result is to find the asymptotic behavior of $E(n, p)$ with accuracy up to positive constants for a broad range of functions $p(n) \to 0$. This is not merely a mathematical exercise; in particular, PCR techniques are permitting screening of ever larger cDNA libraries for ever longer sequences whose probability of occurring in a randomly selected library item is becoming correspondingly ever lower, thereby generating increased practical interest in limiting behavior as $n \to \infty$ and $p(n) \to 0$. To determine the asymptotic behavior of $E(n, p)$, we employ random selection to obtain an upper bound to $\tilde{u}(A, p)$ and also use an explicit construction in Section III. In Section IV, we give a lower bound to $\tilde{u}(A, p)$ which allows us to improve upon the information-theoretic bound (6) when $p \le c\frac{\ln n}{n}$ with any constant $c > 0$. Precise statements concerning the asymptotic behavior of $E(n, p)$ as $n \to \infty$ and $p \to 0$ appear in Section V. The following special cases convey the flavor of our asymptotic results:

$$E(n, p) \sim 2 \qquad \text{if } p \sim \frac{1}{n^2} \tag{16}$$

$$\log_2(\ln n) \lesssim E(n, p) \lesssim 2\sqrt{\ln n}, \qquad \text{if } p \sim \frac{\ln n}{n^2} \tag{17}$$

$$\frac{\log_2 e}{2} \ln n \lesssim E(n, p) \lesssim 2 \ln n, \qquad \text{if } p \sim n^{-\frac{3}{2}} \tag{18}$$

$$\frac{\log_2 e}{4} \frac{(\ln n)^2}{\ln \ln n} \lesssim E(n, p) \lesssim \frac{(\ln n)^2}{\ln \ln n}, \qquad \text{if } p \sim \frac{1}{n} \tag{19}$$

$$\frac{e \log_2 e}{2} (\ln n)^2 \lesssim E(n, p) \lesssim e^2 (\ln n)^2, \quad \text{if } p \sim \frac{\ln n}{n} \tag{20}$$

$$\frac{\log_2 e}{2} n^{\frac{1}{2}} \ln n \lesssim E(n, p) \lesssim 4n^{\frac{1}{2}} \ln n, \qquad \text{if } p \sim n^{-\frac{1}{2}}. \tag{21}$$

Furthermore, for $p(n) = n^{-\beta + o(1)}$, where $0 < \beta < 1$, we show that the asymptotic behavior of $E(n, p)$, up to a positive constant, is $n^{1-\beta} \ln n$. (In the special case $\beta = \frac{1}{2}$, upper and lower bounds on said constant are given in (21).) Since the information-theoretic lower bound (6) also is $O(n^{1-\beta} \ln n)$, the asymptotic efficiency of two-stage algorithms cannot be essentially improved upon when $p(n) = n^{\beta + o(1)}$ even if one were to broaden the domain of algorithms under consideration to include all adaptive schemes with arbitrarily many stages.

III. Upper Bounds to the Expected Number of Unresolved Items

To bound $\tilde{u}(A, p)$ from above for an arbitrary $m \times n$ matrix A and the Bernoulli-p scheme, we use the upper estimate in (12). Although $\overline{X}$ consists of all unresolved and positive items, the number of positive items is small; e.g., it is known (see [5]) not to exceed $\frac{m}{e(1-p)}$. Therefore, to obtain sufficiently good upper bounds to $\tilde{u}(A, p)$ and $E(n, p)$ one can use (12) to write

$$\tilde{u}(A, p) \le \sum_{X \in H_2^n} |\overline{X}| P(X) = \sum_{j \in N_n} \sum_{B_j \subseteq XA^T} P(X). \tag{22}$$

(Here $B_j \subseteq XA^T$ means that the set of active items of the column B_j of the matrix A belongs to the set of active items of the syndrome XA^T for X). Continuing, we consider a probabilistic method to prove the existence of $m \times n$ matrices A with sufficiently small values $\tilde{u}(A, p)$. Specifically, we choose their binary entries randomly and independently to be one with probability s, $0 < s < 1$, and to be zero with probability $1 - s$. This approach was also used in [13] to estimate from above the mean value of inactive unresolved items, but we take all unresolved items into account.

Lemma 1: For any p and s, $0 < p, s < 1$, there exists an $m \times n$ matrix A such that

$$\tilde{u}(A, p) \le n e^{-sm + np(e^{s^2 m} - 1)} + \min\{np, mnpse^{-nps}\}. \tag{23}$$

Proof: We use (22) and estimate the mean value M of

$$\sum_{j \in N_n} \sum_{B_j \subseteq XA^T} P(X) = \sum_{j=1}^{n} \sum_{i=0}^{n} \sum_{X \in H_2^{n,i},\, B_j \subseteq XA^T} P(X)$$

over the class of 2^{mn} matrices whose entries are chosen with respect to the probabilistic scheme above. For fixed $j \in N_n$ and X belonging to the set $H_2^{n,i}$ of all binary n-vectors of Hamming weight i, the event $B_j \subseteq XA^T$ occurs for certain if $i \ge 1$ and $j \in X$, whereas if $j \notin X$ then $B_j \subseteq XA^T$ occurs if and only if none of the m columns of the randomly chosen matrix A^T has a 1 in row j and a 0 in each of the i rows indexed by the 1-entries of X. Hence, $P(B_j \subseteq XA^T)$ equals 1 if $i \ge 1$ and $j \in X$ and equals $(1 - s(1-s)^i)^m$ if $j \notin X$. It follows that

$$\begin{aligned} M &= \sum_{i=0}^{n} (n-i) P(i) (1 - s(1-s)^i)^m + \sum_{i=0}^{n} i P(i) \\ &= n \sum_{i=0}^{n} P(i) (1 - s(1-s)^i)^m \qquad (24) \\ &\quad + \sum_{i=0}^{n} i P(i) (1 - (1 - s(1-s)^i)^m). \qquad (25) \end{aligned}$$

where

$$P(i) = \sum_{X \in H_2^{n,i}} P(X) = \binom{n}{i} p^i (1-p)^{n-i}.$$

Since $(1-s)^i \ge 1 - is$ for any $i = 0, 1, \ldots$, and s, $0 < s < 1$, we have

$$(1 - s(1-s)^i)^m \le (1 - s(1 - is))^m \le e^{-s(1-is)m}.$$

Therefore, (24) does not exceed

$$n\sum_{i=0}^{n}\binom{n}{i}p^i(1-p)^{n-i}e^{-s(1-is)m}$$
$$= ne^{-sm}(1-p)^n\sum_{i=0}^{n}\binom{n}{i}\left(\frac{pe^{s^2m}}{1-p}\right)^i$$
$$= ne^{-sm}\left(1-p+pe^{s^2m}\right)^n \le ne^{-sm+np(e^{s^2m}-1)}.$$

Since

$$(1-s(1-s)^i)^m > 0$$

and

$$(1-s(1-s)^i)^m \ge 1-ms(1-s)^i$$

we can upper-bound (25) both by np and by

$$mns\sum_{i=1}^{n}\binom{n-1}{i-1}p^i(1-p)^{n-i}(1-s)^i$$
$$= mnps(1-s)\sum_{j=0}^{n-1}\binom{n-1}{j}p^j(1-p)^{n-1-j}(1-s)^j$$
$$= mnps(1-s)(1-p)^{n-1}\sum_{j=0}^{n-1}\binom{n-1}{j}\left(\frac{p(1-s)}{1-p}\right)^j$$
$$= mnps(1-s)(1-ps)^{n-1} \le mnpse^{-nps}.$$

This completes the proof since there exists a matrix A for which $\tilde{u}(A,\,p)$ does not exceed the mean M of this value over the class of matrices under consideration. □

One easily checks that $mnpse^{-nps} \le \frac{m}{e}$ and, for some values of parameters, $\frac{m}{e} \le np$ holds. However, we shall see that $np = o(m)$ as $n \to \infty$ for the number m of rows of matrices A which minimize $E(A,\,p)$.

Now we give asymptotic estimates which follow from Lemma 1 and show how the asymptotic behavior of $E(n,\,p)$ depends on the function $p = p(n)$. For a constant $c > 0$, consider the function $f_1(x) = 1 + x - x\ln\frac{x}{c} - c$ and note that $f_1(x)$ decreases for increasing $x \ge c$ from $f_1(c) = 1$ to $-\infty$; hence, for $x > c$, $f_1(x)$ has a unique zero which will be denoted by $\mu_1 = \mu_1(c)$.

Theorem 1: Let $c > 0$, $\mu_1 = \mu_1(c)$, and $n \to \infty$. Then

$$E(n,\,p) \lesssim \frac{(\ln n)^2}{\ln\frac{\ln n}{np}}, \qquad \text{if } p \ge \frac{\ln n}{n^2} \text{ and } \frac{pn}{\ln n} \to 0 \qquad (26)$$

$$E(n,\,p) \lesssim \left(\mu_1^2\ln\frac{\mu_1}{c}\right)(\ln n)^2, \qquad \text{if } p = c\frac{\ln n}{n} \qquad (27)$$

$$E(n,\,p) \lesssim 4np\ln n, \qquad \text{if } \frac{pn}{\ln n} \to \infty. \qquad (28)$$

Proof: By Lemma 1, for any $s = \frac{1}{k}$ with $k > 1$ there exists an $m\times n$ matrix A of any size $m\times n$ such that $E(A,\,p) = m + \tilde{u}(A,\,p)$ where

$$\tilde{u}(A,\,p) \le \exp\left\{-\frac{m}{k} + \ln n + np\left(e^{\frac{m}{k^2}} - 1\right)\right\} + np. \qquad (29)$$

Therefore, to prove that $E(n,\,p) \lesssim m(n)$ where $m(n) \to \infty$ as $n \to \infty$ it is sufficient to show that for $m = \lceil(1+4\varepsilon)m(n)\rceil$, with $\varepsilon > 0$ as small as one wishes, and a certain choice of $k = k(n) > 1$ we have

$$np = o(m(n)) \qquad (30)$$

and

$$-\frac{m}{k} + \ln n + np\left(e^{\frac{m}{k^2}} - 1\right) \le (-\varepsilon + o(1))\ln n.$$

For a suitable choice of $k = k(n)$, we use the fact that the minimum over m of the exponent in (29) is attained when

$$m = k^2\ln\frac{k}{np}, \qquad \text{if } k > np. \qquad (31)$$

For the case in which $p \ge \frac{\ln n}{n^2}$ and $\frac{pn}{\ln n} \to 0$, we put

$$m = \left\lceil(1+4\varepsilon)\frac{(\ln n)^2}{\ln\frac{\ln n}{np}}\right\rceil$$

and define k by the condition

$$m = (1-\varepsilon)k^2\ln\frac{\ln n}{np}.$$

Note that $k > 1$, since otherwise, for any $\varepsilon > 0$

$$(1+4\varepsilon)\frac{(\ln n)^2}{\ln\frac{\ln n}{np}} \le m \le (1-\varepsilon)\ln\frac{\ln n}{np}$$

which contradicts the condition $p \ge \frac{\ln n}{n^2}$. We can assume that ε is sufficiently small (any $\varepsilon \le \frac{1}{5}$ will suffice) that $(1+4\varepsilon)(1-\varepsilon) \ge (1+\varepsilon)^2$. Then

$$m^2 \ge m(1+4\varepsilon)(\ln n)^2 \Big/ \left(\ln\frac{\ln n}{np}\right)$$
$$= (1+4\varepsilon)(1-\varepsilon)k^2(\ln n)^2 \ge (1+\varepsilon)^2k^2(\ln n)^2$$

so $m \ge (1+\varepsilon)k\ln n$. Therefore,

$$-\frac{m}{k} + \ln n + np\left(e^{\frac{m}{k^2}} - 1\right)$$
$$\le -(1+\varepsilon)\ln n + \ln n + np\left(\frac{\ln n}{np}\right)^{1-\varepsilon}$$
$$= \left(-\varepsilon + \left(\frac{np}{\ln n}\right)^{\varepsilon}\right)\ln n = (-\varepsilon + o(1))\ln n.$$

Since $\frac{\ln n}{pn} \to \infty$, we have $\ln\frac{\ln n}{pn} = o(\frac{\ln n}{pn})$. Hence $np = o(\frac{m}{\ln n})$ and (30) holds, which establishes (26).

Moving to the case $p = c\frac{\ln n}{n}$, $c > 0$, we put

$$m = \left\lfloor(1+\varepsilon)^2(\ln n)^2\mu_1^2\ln\frac{\mu_1}{c}\right\rfloor$$

and consider

$$k = (1+\varepsilon)\mu_1\ln n$$

where $\mu_1 = \mu_1(c)$. Note that $m + 1 \ge (1+\varepsilon)(\ln n)k\mu_1\ln\frac{\mu_1}{c}$ and $m \le k^2\ln\frac{\mu_1}{c}$. Therefore, using the definition of μ_1 we have

$$-\frac{m}{k} + \ln n + np\left(e^{\frac{m}{k^2}} - 1\right)$$
$$\le \frac{1}{k} + \left(-(1+\varepsilon)\mu_1\ln\frac{\mu_1}{c} + 1 + \mu_1 - c\right)\ln n$$
$$\le \left(-\varepsilon\mu_1\ln\frac{\mu_1}{c} + o(1)\right)\ln n \le (-\varepsilon + o(1))\ln n.$$

In this case, we have $np = O(\frac{m}{\ln n})$ and hence (30) holds, establishing (27).

Finally, in the case $\frac{pn}{\ln n} \to \infty$ we put

$$m = \lfloor 4(1+\varepsilon)np \ln n \rfloor \quad \text{and} \quad k = 2np.$$

Then, $\frac{m+1}{k} \ge 2(1+\varepsilon)\ln n$ and $\frac{m}{k^2} \le (1+\varepsilon)\frac{\ln n}{np}$, so $\frac{m}{k^2} \to 0$. Using $e^x - 1 = x(1+o(1))$ for $x \to 0$, we get

$$\begin{aligned} -\frac{m}{k} + \ln n + pn\left(e^{\frac{m}{k^2}} - 1\right) &\le \frac{1}{k} + (-2(1+\varepsilon) + 1 \\ &\quad + (1+\varepsilon)(1+o(1)))\ln n \\ &= (-\varepsilon + o(1))\ln n. \end{aligned}$$

In this case, we also have $np = O(\frac{m}{\ln n})$, so again (30) holds, which establishes (28) and completes the proof. □

In particular, for $p = \frac{1}{n}$, (26) improves upon [13, Theorem 4.1] by a factor of e. It should be noted that, for the functions $p = p(n)$ considered, we take into account all unresolved items, not just inactive unresolved ones as in [13]. Although the number of active unresolved items does not exceed np, this value tends to infinity in the last two cases and also in the first case provided p is not too small.

In Theorem 1, the restriction $p \ge \frac{\ln n}{n^2}$ reflects the requirement that the probability s in Lemma 1 must be less than 1. Now we give a simple explicit construction which is valid for all p and better than random selection for sufficiently small p.

Lemma 2: For any n, m, and p, $1 \le m \le n$, $0 < p < 1$

$$E(n, p) \le m + pn\left\lceil \frac{n}{m} \right\rceil. \tag{32}$$

Proof: Put $h = \lceil \frac{n}{m} \rceil$, $k = \lceil \frac{n}{h} \rceil$, and note that $h \ge \frac{n}{m}$ and $\frac{n}{h} \le k \le \frac{n+h-1}{h}$. Hence,

$$k \le m \quad \text{and} \quad (k-1)h + 1 \le n \le kh. \tag{33}$$

Consider the matrix $A = A(n, m)$ with k rows given by specifying the subsets of N_n that comprise their active items as follows:

$$A_i = \{(i-1)h+1, \ldots, (i-1)h+h\}, \qquad i = 1, \ldots, k-1$$

and

$$A_k = \{(k-1)h+1, \ldots, n\}.$$

(The matrix $A(11, 5)$ with $h = 3$ and $k = 4$ is given in Example 2.) From (33), it follows that the number of rows of A does not exceed m and that the number h_0 of active items of row k satisfies the inequality $1 \le h_0 \le h$. It is clear that an item $j \in A_i$ is unresolved if and only if $|A_i| \ge 2$ and at least one of the active items of A_i is positive. Hence, using $\sum_{i=1}^{k} |A_i| = n$, we have

$$\begin{aligned} \tilde{u}(A, p) &= \sum_{i:\, |A_i| \ge 2} |A_i| \left(1 - (1-p)^{|A_i|}\right) \\ &\le \sum_{i:\, |A_i| \ge 2} |A_i| \left(1 - (1-p)^h\right) \le pnh. \end{aligned}$$

Lemma 2 is proved. □

Example 2: The matrix $A(11, 5)$ has the following form:

$$\begin{array}{ccccccccccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1. \end{array}$$

We remark that this explicit construction in fact achieves the asymptotically optimal form of $E(n, p)$ in cases in which $p(n)$ decays to 0 sufficiently rapidly as $n \to \infty$, e.g., when $p(n) \sim n^{-2}$. Toward verifying this, we first establish that for $n \ge 2$

$$E(n, p) \ge \min\left(2,\, 1 + n\left(1 - (1-p)^n\right)\right).$$

Indeed, let $E(n, p)$ be attained at an $m \times n$ matrix A. If $m \ge 2$, then $E(n, p) \ge 2$, so it remains to consider only the case when this A consists of one row. If this row has $h \ge 2$ active items (the case of $h = 1$ is trivial), then $E(A, p) = 1 + n - h(1-p)^h$. This function is unimodal in h viewed as a positive real variable, assuming a minimum of $\frac{1}{e \ln(1-p)}$ at $x = -\frac{1}{\ln(1-p)}$. Since $h \le n$, it follows that, if $n \le -\frac{1}{\ln(1-p)}$, then the integer h that minimizes $1 + n - h(1-p)^h$ is $h = n$, giving $E(A, p) \ge 1 + n(1 - (1-p)^n)$. This leaves only the case $n \ge -\frac{1}{\ln(1-p)}$, in which we have

$$E(A, p) \ge 1 + n + \frac{1}{e \ln(1-p)} \ge 1 + n - \frac{n}{e} > 2$$

for $n \ge 2$, so the inequality is established. This, in turn, gives the following corollary to the effect that (32) with $m = 1$ is asymptotically tight if $p(n) \sim \frac{1}{n^2}$.

Corollary 1: If $n \to \infty$, then

$$E(n, p) \sim 2 \quad \text{when } p = p(n) \sim \frac{1}{n^2}. \tag{34}$$

Setting $m = \lceil n\sqrt{p} \rceil$ in (32) so that $m < n\sqrt{p} + 1$ and $\lceil \frac{n}{m} \rceil < \frac{1}{\sqrt{p}} + 1$, we get the following statement.

Corollary 2: For any n and p, $0 < p < 1$

$$E(n, p) < 1 + 2n\sqrt{p} + np. \tag{35}$$

Corollary 2 implies that $E(n, p) \lesssim 2c + 1$ if $pn^2 \to c^2$ as $n \to \infty$ where $c \ge 0$. For $p = (c\frac{\ln n}{n})^2$, (35) gives $E(n, p) \lesssim 2c \ln n$ which improves upon (26) for $c < \frac{1}{2}$. In fact, (35) also improves on (26) for any smaller p that meets the conditions of (26). In particular, for $p = (\frac{c}{n})^2 \ln n$, (35) shows that $E(n, p) \lesssim 2c\sqrt{\ln n}$, while (26) applies only for $c \ge 1$ where it gives the weaker result $E(n, p) \lesssim \ln n$.

IV. Universal Bounds to the Expected Number of Unresolved Items

By universal bounds we shall mean bounds valid for all (two-stage) reconstruction algorithms. In particular, the information-theoretic bound (6) is universal. Another universal bound can be obtained with the help of averaging and linear programming; this approach was introduced by Knill [13].

We fix a probability distribution P on the set H_2^n and a binary matrix A of size $m \times n$. For any subset $X \subseteq N_n$ of active items, we have defined $\overline{X} = X(S)$ using the syndrome $S = XA^T$ (see (11)). By (12) and (14) we have

$$\tilde{u}(A, P) \geq \sum_{X \in H_2^n} P(X)\left(|\overline{X}| - |X|\right).$$

A function $F\colon H_2^n \to H_2^n$ is called a *covering* operator on H_2^n if $X \subseteq F(X)$ for any $X \in H_2^n$. Using averaging and linear programming, one of the authors recently [15] proved that for any covering operator F on H_2^n and any $i = 1, \ldots, n$

$$\frac{1}{\binom{n}{i}} \sum_{X \in H_2^{n,\, i}} |F(X)| \geq \frac{n}{\sqrt[i]{M_i}}$$

where $M_i = |\{F(X)\colon X \in H_2^{n,\, i}\}|$. Note that $F(X) = \overline{X}$ is a covering operator on $H_2^n = \bigcup_{i=0}^{n} H_2^{n,\, i}$ and that for any $m \times n$ matrix A and any i, $M_i \leq 2^m$ when $F(X) = \overline{X}$. Therefore, we get the following statement.

Theorem 2: For any symmetric distribution P and matrix A of size $m \times n$

$$\tilde{u}(A, P) \geq n \sum_{i=1}^{n} P(i) 2^{-\frac{m}{i}} - \sum_{i=1}^{n} iP(i) \tag{37}$$

where $P(i) = \sum_{X \in H_2^{n,\, i}} P(X)$, and for any integer k, $1 \leq k \leq n$

$$\tilde{u}(A, P) \geq P(k)\left(n2^{-\frac{m}{k}} - k\right). \tag{38}$$

The inequality (38) is a minor improvement of [13, Theorem 3.2]

$$\tilde{u}(A, P) \geq P(k)\left(n2^{-\frac{m}{k}}\left(1-\frac{4}{k}\right) - k\right). \tag{39}$$

However, as distinguished from (39), one can use (38) when the maximum member in (37) is obtained for $i = k \leq 4$, as we do in the following. Note also that the ratio of two successive members of the sum in (37) shows that for finding this maximizing k one can use the asymptotic equality

$$\frac{P(k+1)}{P(k)} \sim 2^{-\frac{m}{k(k+1)}}. \tag{40}$$

Now we give asymptotic results for the Bernoulli p-scheme when $p \to 0$ as $n \to \infty$. Note that in this case

$$P(i) = \binom{n}{i} p^i (1-p)^{n-i}, \qquad \sum_{i=1}^{n} iP(i) = np$$

and (40) means that

$$k \sim np2^{\frac{m}{k^2}}, \qquad \text{if } k > np,\ \ k = o(n),\ \ \text{and}\ \ k \to \infty$$

(compare with (31)). From (38) and the definition (15) it follows that for any integer k, $1 \leq k \leq n$

$$E(n, p) \geq \min_{1 \leq m \leq n} \left(m + P(k)(n2^{-\frac{m}{k}} - k)\right). \tag{41}$$

Note that this remains valid if we consider the minimum over all (not necessarily integer) numbers m, $1 \leq m \leq n$.

As an example, in the case $k = 1$ we have

$$E(n, p) \geq \min_{1 \leq m \leq n} \left(m + np(1-p)^{n-1}(n2^{-m} - 1)\right).$$

The dominant portion of the quantity to be minimized here is $m + n^2 p 2^{-m}$. Considering separately the case in which the minimum is obtained at $m = 1$ and the case in which differentiation with respect to m reveals that the minimum occurs at $m = \log_2(n^2 p \ln 2)$, we get the following statement.

Corollary 3: If $pn = o(1)$ as $n \to \infty$, then

$$E(n, p) \gtrsim \begin{cases} 1 + \frac{1}{2}\, n^2 p, & \text{if } n^2 p \leq 2\log_2 e \\ \log_2\left(\frac{e}{\log_2 e}\, n^2 p\right), & \text{if } n^2 p \geq 2\log_2 e. \end{cases} \tag{42}$$

In particular, (42) implies that if $\frac{\ln pn}{\ln n} \to -\alpha$ (i.e., if $p = n^{-1-\alpha+o(1)}$), where $0 < \alpha < 1$, then

$$E(n, p) \gtrsim (1-\alpha)\log_2 n. \tag{43}$$

Now we shall see that this bound can be strengthened for $0 < \alpha \leq \frac{1}{4}$ by proper selection of $k \geq 2$.

For a constant $c > 0$ note that the function

$$f_2(x) = x - 2x\ln\frac{x}{c} - c + 1$$

decreases for increasing $x \geq c$ from $f(c) = 1$ to $-\infty$. Hence, for $x > c$, $f_2(x)$ has a unique zero which will be denoted by $\mu_2 = \mu_2(c)$. We shall also use the fact that $\mu_2 \ln\frac{\mu 2}{c} < 1$ because $x - x\ln\frac{x}{c} - c$ decreases for increasing $x \geq c$ and hence is negative when $x > c$.

Theorem 3: Let $c > 0$, $\mu_2 = \mu_2(c)$, $0 < \alpha \leq \frac{1}{4}$, and $n \to \infty$. Then

$$E(n, p) \gtrsim \frac{1}{2}\left\lfloor\frac{1}{2\alpha}\right\rfloor \log_2 n, \qquad \text{if } \frac{\ln pn}{\ln n} \to -\alpha \tag{44}$$

$$E(n, p) \gtrsim \frac{\log_2 e}{4}\,\frac{(\ln n)^2}{\ln\frac{\ln n}{np}}, \qquad \text{if } \frac{\ln pn}{\ln n} \to 0,\ \ \frac{pn}{\ln n} \to 0 \tag{45}$$

$$E(n, p) \gtrsim \left(\mu_2^2 \log_2\frac{\mu_2}{c}\right)(\ln n)^2, \qquad \text{if } p = c\,\frac{\ln n}{n}. \tag{46}$$

Proof: Due to (41), to prove that $E(n, p) \gtrsim l(n)$ where $l(n) \to \infty$ as $n \to \infty$, it suffices to show that

$$np = o(l(n)) \tag{47}$$

and there exist integers $k = k(n) \geq 2$ such that the inequality $m \leq (1-\varepsilon)l(n)$ with $\varepsilon > 0$ implies $l(n) = o(nP(k)2^{-\frac{m}{k}})$ or equivalently

$$\left(\ln n + \ln P(k) - \frac{m}{k}\,\ln 2 - \ln l(n)\right) \to \infty. \tag{48}$$

Note that, for any integer $k = k(n) \geq 1$, if $n \to \infty$, $\frac{k}{n} \to 0$, $p \to 0$, and hence $\frac{k-np}{n-k} \to 0$, then

$$P(k) \gtrsim \frac{e^{\frac{-1}{12k}}}{\sqrt{2\pi k}}\left(\frac{k}{np}\right)^{-k}\left(\frac{n(1-p)}{n-k}\right)^{n-k}$$

by the Stirling formula, and

$$\left(1 + \frac{k-np}{n-k}\right)^{n-k} \geq e^{(k-np)(1+o(1))}.$$

Thus, for these asymptotics

$$\ln P(k) \geq -k\ln\frac{k}{np} + (k-np)(1+o(1)) - \frac{1}{2}\,\ln k + O(1). \tag{49}$$

In the case $\frac{\ln pn}{\ln n} \to -\alpha$ (or $p = n^{-1-\alpha+o(1)}$), $0 < \alpha \leq \frac{1}{4}$, we put $l(n) = \frac{1}{2}\lfloor\frac{1}{2\alpha}\rfloor\log_2 n$ and $k(n) = \lfloor\frac{1}{2\alpha}\rfloor$. Then

$np = o(l(n))$, $k(n)$ does not depend on n, $k(n) \geq 2$, and $\ln P(k) \geq (-k\alpha + o(1)) \ln n$, by (49). Therefore, for $m \leq (1-\varepsilon)l(n)$ we have

$$\ln\left(nP(k)2^{-\frac{m}{k}}\right) \geq \left(1 - k\alpha - \frac{1}{2}(1-\varepsilon) + o(1)\right) \ln n$$
$$\geq \left(\frac{\varepsilon}{2} + o(1)\right) \ln n.$$

In the second and third cases, we denote the right-hand sides of (45) and (46) by $l(n)$ and put $k = \lceil h \rceil$ where, respectively,

$$h = \frac{\ln n}{2 \ln \frac{\ln n}{np}} \quad \text{and} \quad h = \mu_2 \ln n.$$

Next we check that in both cases h and k tend to infinity and (47) holds. It is clear for the third case. In the second case, $\ln \frac{\ln n}{np} = o(\ln n)$ since $\frac{\ln pn}{\ln n} \to 0$; hence, $h \to \infty$ and $\ln n = o(l(n))$ while $pn = o(\ln n)$. Note that in both cases

$$l(n) = Ch \log_2 n, \qquad \text{with } 0 < C < 1,$$

where, respectively,

$$C = \frac{1}{2} \quad \text{and} \quad C = \mu_2 \ln \frac{\mu_2}{c}.$$

Therefore, for $m \leq (1-\varepsilon)l(n)$, we obtain

$$n2^{-\frac{m}{k}} \geq n2^{-\frac{m}{h}} \geq n^{1-C+C\varepsilon}. \tag{50}$$

Now we show that if $k \to \infty$ we, in fact, can replace $k = \lceil h \rceil$ in (49) by h through $k - 1 < h \leq k$. Since $\ln \frac{k}{h} \leq \frac{k-h}{h} < \frac{1}{k-1}$ and

$$-k \ln \frac{k}{pn} = -k \ln \frac{h}{pn} - k \ln \frac{k}{h}$$

we have

$$-k \ln \frac{k}{pn} \geq -(1 + o(1))h \ln \frac{h}{pn} + O(1). \tag{51}$$

In the case when $\frac{pn}{\ln n} \to 0$ and $h = o(\ln n)$, we have $k = o(\ln n)$ and

$$-h \ln \frac{h}{pn} \geq -h \ln \frac{\ln n}{pn} = -\frac{1}{2} \ln n.$$

Therefore, using (49)–(51) we get

$$\ln\left(P(k)n2^{-\frac{m}{k}}\right) \geq \left(\frac{\varepsilon}{2} + o(1)\right) \ln n$$

while $l(n)$ grows slower than $(\ln n)^2$. In the case when $p = c\frac{\ln n}{n}$ using (49)–(51) we get

$$\ln\left(P(k)n2^{-\frac{m}{k}}\right)$$
$$\geq \left(-\mu_2 \ln \frac{\mu_2}{c} + \mu_2 - c + 1 - (1-\varepsilon)\mu_2 \ln \frac{\mu_2}{c} + o(1)\right) \ln n$$
$$\geq \left(\varepsilon \mu_2 \ln \frac{\mu_2}{c} + o(1)\right) \ln n$$

while $l(n)$ grows as $(\ln n)^2$. □

In conclusion, we verify that all asymptotic lower bounds of Corollary 3 and Theorem 3 are better than the information-theoretic asymptotic lower bound

$$E(n, p) \gtrsim np \log_2 \frac{1}{p}, \qquad \text{as } n \to \infty \text{ and } p \to 0 \tag{52}$$

which follows from (6). Indeed, if $\frac{\ln pn}{\ln n} \to -\alpha, 0 \leq \alpha \leq 1$, then $p = n^{-1-\alpha+o(1)}$ and the right-hand side of (52) is asymptotically equal to $(1+\alpha)np \log_2 n$. This is only $o(1)$ when $\alpha > 0$. If $\alpha = 0$ and $\frac{pn}{\ln n} \to 0$, then $\frac{\ln n}{pn} \to \infty$, $\ln \frac{\ln n}{pn} = o(\frac{\ln n}{pn})$, and, hence, $pn = o(\frac{\ln n}{\ln \frac{\ln n}{pn}})$. This means that (45) is also asymptotically better than (52). In the case $p = c\frac{\ln n}{n}$, (52) takes the form $E(n, p) \gtrsim c \log_2 e (\ln n)^2$ and has the same order of magnitude as (46). However, one can verify that (46) is still asymptotically better since

$$\mu_2^2 \ln \frac{\mu_2}{c} > c, \qquad \text{for all } c > 0. \tag{53}$$

Indeed, by the definition of $f_2(x)$ and $\mu_2 = \mu_2(c)$, (53) is equivalent to $\mu_2^2 - (c-1)\mu_2 - 2c > 0$. The last inequality is satisfied if $\mu_2 > \nu$ where $\nu = \nu(c)$ is the largest root of the quadratic equation $x^2 - (c-1)x - 2c = 0$ $(\nu > c)$. Since $f_2(x)$ decreases for increasing $x > c$, $f_2(c) = 1$, and $f_2(\mu_2) = 0$, we have $\mu_2 > \nu$ if $f_2(\nu) > 0$. Using $c = \frac{\nu(\nu+1)}{\nu+2}$ we get

$$f_2(\nu) = \nu - 2\nu \ln \frac{\nu}{c} - c + 1$$
$$> \nu - 2\nu\left(\frac{\nu}{c} - 1\right) - c + 1$$
$$= \nu - \frac{2\nu}{\nu+1} - \frac{\nu(\nu+1)}{\nu+2} + 1 = \frac{2}{(\nu+1)(\nu+2)} > 0.$$

Thus, the information-theoretic bound (52) is weaker than the bounds (44)–(46). However, the information-theoretic bound does give the proper asymptotic behavior when $\frac{\ln p}{\ln n} \to -\beta$, $0 < \beta < 1$; see (60).

V. Asymptotic Behavior of the Minimum Expected Number of Tests

Now we summarize results on asymptotic behavior of $E(n, p)$ which follow from the upper and lower bounds proved in the previous sections.

First we note an inference which follows from the information-theoretic bound (6) and Corollary 2.

Corollary 4: If $n \to \infty$, then

$$E(n, p) = o(n), \qquad \text{if and only if } p = p(n) \to 0. \tag{54}$$

From Corollary 2 it also follows that

$$E(n, p) \sim 1, \qquad \text{if } p = o\left(\frac{1}{n^2}\right). \tag{55}$$

In the case $pn^2 \to c$, $c > 0$, the value $E(n, p)$ is restricted from above and below by constants which can be defined from Corollaries 2 and 3. In the case when $pn^2 \to \infty$ and $\frac{\ln pn}{\ln n} \to -1$ (i.e., $p = n^{-2+o(1)}$) these corollaries give

$$2 \log_2 (n\sqrt{p}) \lesssim E(n, p) \lesssim 2n\sqrt{p} \tag{56}$$

and do not allow us to determine the order of magnitude of $E(n, p)$. However, now we show that Theorem 1, Theorem 3, and (52) allow us to determine it outside the range considered when $\frac{\ln pn}{\ln n} \to \gamma$ (i.e., $p = n^{-1+\gamma+o(1)}$) where $-1 < \gamma < 1$.

Theorem 4: Let $c > 0, \mu_1 = \mu_1(c), \mu_2 = \mu_2(c), 0 < \alpha < 1$, $c(\alpha) = 1 - \alpha$ if $\frac{1}{4} < \alpha < 1$, $c(\alpha) = \frac{1}{2}\lfloor \frac{1}{2\alpha} \rfloor$ if $0 < \alpha \leq \frac{1}{4}$, $0 < \beta < 1$, and $n \to \infty$. Then

$$c(\alpha) \log_2 n \lesssim E(n, p) \lesssim \frac{1}{\alpha} \ln n, \qquad \text{if } \frac{\ln pn}{\ln n} \to -\alpha \tag{57}$$

$$\frac{\log_2 e}{4}\,\frac{(\ln n)^2}{\ln\frac{\ln n}{np}} \lesssim E(n,\,p) \lesssim \frac{(\ln n)^2}{\ln\frac{\ln n}{np}},$$

$$\text{if } \frac{\ln pn}{\ln n} \to 0 \quad \text{and} \quad \frac{pn}{\ln n} \to 0 \tag{58}$$

$$\left(\mu_2^2 \log_2 \frac{\mu_2}{c}\right)(\ln n)^2 \lesssim E(n,\,p) \lesssim \left(\mu_1^2 \ln \frac{\mu_1}{c}\right)(\ln n)^2,$$

$$\text{if } p = c\frac{\ln n}{n} \tag{59}$$

$$\beta np \log_2 n \lesssim E(n,\,p) \lesssim 4np\ln n, \qquad \text{if } \frac{\ln p}{\ln n} \to -\beta. \tag{60}$$

Proof: Since $\frac{\ln pn}{\ln n} \to -\alpha$ where $0 < \alpha < 1$ implies that $\ln \frac{\ln n}{np} \sim \alpha \ln n$, the upper asymptotic bound in (57) follows from (26) and the upper asymptotic bounds in (58) and (59) coincide with those in (26) and (27). The lower asymptotic bounds in (57)–(59) coincide, respectively, with those in (43)–(46). Finally, the condition $\frac{\ln p}{\ln n} \to -\beta$ where $0 < \beta < 1$ implies that $\frac{pn}{\ln n} \to \infty$ and $\frac{1}{p} \sim \beta \ln n$. Therefore, the upper asymptotic bound in (60) follows from (28) and the lower one follows from (52). □

It is significant that the lower asymptotic bound in (60) was obtained by using the information-theoretic bound and hence is valid for all adaptive reconstruction algorithms. This means that, for $p(n) = n^{-\beta+o(1)}$ where $0 < \beta < 1$, the optimal efficiency of such algorithms (up to a positive constant) can be reached in the class of two-stage algorithms.

The important special cases (16)–(21) presented in Section II follow from (34), (56)–(60) by noting that $\mu_1(1) = e$ and $\mu_2(1) = \sqrt{e}$. We earlier noted that the bounds (19) were obtained by Knill [13] with the extra factor of e in the upper bound. By (58) they remain valid for any function $p = p(n) = \frac{(\ln n)^{o(1)}}{n}$ (we have $p = \frac{c}{n}$ if $o(1) = \frac{\ln c}{\ln \ln n}$).

An interesting open problem is to strengthen the bounds (56). One can expect that the upper bound in (56) based on the explicit construction of Lemma 2 and Corollary 2 is asymptotically tight for $\frac{1}{n^2} \leq p \leq (\frac{\ln n}{2n})^2$.

References

[1] M. Aigner, *Combinatorial Search*. Stuttgart, Germany: Wiley–Teubner, 1988.

[2] E. Barillot, B. Lacroix, and D. Cohen, “Theoretical analysis of library screening using an N-dimensional pooling strategy,” *Nucleic Acids Res.*, vol. 19, pp. 6241–6247, 1991.

[3] T. Berger, N. Mehravari, D. Towsley, and J. Wolf, “Random multiple-access communication and group testing,” *IEEE Trans. Commun.*, vol. COM-32, pp. 769–779, 1984.

[4] T. Berger, J. W. Mandell, and P. Subrahmanya, “Maximally efficient two-stage group testing,” *Biometrics*, vol. 56, pp. 107–114, Sept. 2000.

[5] T. Berger and V. I. Levenshtein, “Application of cover-free codes and combinatorial block-designs to two-stage testing,” *Discr. Appl. Math.*, submitted for publication.

[6] A. G. Djachkov and V. V. Rykov, “A survey of superimposed distance codes,” *Probl. Contr. and Inform. Theory*, vol. 12, pp. 11–13, 1983.

[7] R. Dorfman, “The detection of defective members of large populations,” *Ann. Math. Statist.*, vol. 14, pp. 436–440, 1943.

[8] D. Z. Du and F. K. Hwang, *Combinatorial Group Testing and Its Applications*, Singapore: World Scientific, 1993.

[9] P. Erdos and A. Renyi, “On two problems of information theory,” *Publ. Math. Ins. Hung. Acad. Sci.*, vol. 8, pp. 241–254, 1963.

[10] W. T. Federer, *Statistics and Society: Data Collection and Interpretation*. New York: Marcel Dekker, 1991.

[11] F. K. Hwang and V. T. Soś, “Non-adaptive hypergeometric group testing,” *Stud. Sci. Math. Hung.*, vol. 22, pp. 257–263, 1987.

[12] W. H. Kautz and R. R. Singleton, “Nonrandom binary superimposed codes,” *IEEE Trans. Inform. Theory*, vol. IT-10, pp. 363–377, Dec. 1964.

[13] E. Knill, “Lower bounds for identifying subset members with subset queries,” in *Proc. 6th Annu. ACM-SIAM Symp. Discrete Algorithms* San Francisco, CA, Jan. 22–24, 1995, ch. 40, pp. 369–377.

[14] V. I. Levenshtein, “Efficient reconstruction of sequences,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 2–22, Jan. 2001.

[15] ——, “A universal bound for a covering in regular posets and its application to pool testing,” *Discr. Math.*, submitted for publication.

[16] C. H. Li, “A sequential method for screening experimental variables,” *J. Amer. Statist. Assoc.*, vol. 57, pp. 455–477, 1962.

[17] D. Kurtz and M. Sidi, “Multiple-access algorithms via group testing for heterogeneous population of users,” *IEEE Trans. Commun.*, vol. 36, pp. 1316–1323, Dec. 1988.

[18] A. Macula, “Probabilistic nonadaptive group testing in the presence of errors and DNA library screening,” *Ann. Combin.*, vol. 3, pp. 61–69, 1999.

[19] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.*. Amsterdam, The Netherlands: North-Holland, 1977.

[20] Q. A. Nguyen and T. Zeisel, “Bounds on constant weight binary superimposed codes,” *Probl. Contr. and Inform. Theory*, vol. 17, pp. 223–230, 1988.

[21] C. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.*, vol. 27, pp. 379-423 and 623–656, 1948.

[22] M. Sobel and P. A. Groll, “Group testing to eliminate efficiently all defectives in a binomial sample,” *Bell Syst. Tech. J.*, vol. 38, pp. 1179–1252, 1959.

[23] J. Wolf, “Born again group testing: Multiaccess communications,” *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 185–191, Mar. 1985.

Toby Berger (S’60–M’66–SM’74–F’78) was born in New York, NY, on September 4, 1940. He received the B.E. degree in electrical engineering from Yale University, New Haven, CT, in 1962 and the M.S. and Ph.D. degrees in applied mathematics from Harvard University, Cambridge, MA, in 1964 and 1966, respectively.

From 1962 to 1968, he was a Senior Scientist at Raytheon Company, Wayland, MA, specializing in communication theory, information theory, and coherent signal processing. In 1968, he joined the faculty of Cornell University, Ithaca, NY, where he is presently the Irwin and Joan Jacobs Professor of Engineering. His research interests include information theory, random fields, communication networks, wireless communications, video compression, voice and signature compression and verification, infobiology, quantum information theory, and coherent signal processing. He is the author/coauthor of *Rate Distortion Theory: A Mathematical Basis for Data Compression*, *Digital Compression for Multimedia: Principles and Standards*, and *Information Measures for Discrete Random Fields*.

Prof. Berger has served as Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY and as President of the IEEE Information Theory Group. He is a Fellow of the Guggenheim Foundation, the Japan Society for Promotion of Science, the Ministry of Education of the People’s Republic of China, and the Fulbright Foundation. In 1982, he received the Frederick E. Terman Award of the American Society for Engineering Education. The IEEE Information Theory Society has designated him the 2002 Shannon Award Recipient.

Vladimir I. Levenshtein (M’01) was born in Moscow, U.S.S.R., on May 20, 1935. He graduated from the Mathematical Department of Moscow State University in 1958 and received the Candidate of Science degree in physics and mathematics in 1962 from the Institute for Applied Mathematics, Moscow. He received the Doctor of Science degree in physics and mathematics from Moscow State University in 1983.

He has since been with the Keldysh Institute for Applied Mathematics of the Russian Academy of Sciences, Moscow, where he currently holds the position of Leading Scientific Researcher. His main scientific interests are in lexicographic codes, synchronization with finite delay, codes with correction of deletions and insertions (Levenshtein metric), efficient coding of integers, duality in bounding codes and designs, efficient reconstruction of sequences, and algorithms for disjunctive testing.