Group Testing with Probabilistic Tests: Theory, Design and Application

Mahdi Cheraghchi, *Member, IEEE*, Ali Hormati, *Member, IEEE*, Amin Karbasi, *Student Member, IEEE*, and Martin Vetterli *Fellow, IEEE*

Abstract—Identification of defective members of large populations has been widely studied in the statistics community under the name of group testing. It involves grouping subsets of items into different pools and detecting defective members based on the set of test results obtained for each pool.

In a classical noiseless group testing setup, it is assumed that the sampling procedure is fully known to the reconstruction algorithm, in the sense that the existence of a defective member in a pool results in the test outcome of that pool to be positive. However, this may not be always a valid assumption in some cases of interest. In particular, we consider the case where the defective items in a pool can become independently *inactive* with a certain probability. Hence, one may obtain a negative test result in a pool despite containing some defective items. As a result, any sampling and reconstruction method should be able to cope with two different types of uncertainty, i.e., the unknown set of defective items and the partially unknown, probabilistic testing procedure.

In this work, motivated by the application of detecting infected people in viral epidemics, we design non-adaptive sampling procedures that allow successful identification of the defective items through a set of probabilistic tests. Our design requires only a small number of tests to single out the defective items. In particular, for a population of size N and at most K defective items with activation probability p, our results show that $M = O(K^2 \log{(N/K)}/p^3)$ tests is sufficient if the sampling procedure should work for *all* possible sets of defective items, while $M = O(K \log{(N)}/p^3)$ tests is enough to be successful for any single set of defective items. Moreover, we show that the defective members can be recovered using a simple reconstruction algorithm with complexity of $O(MN)$.

Index Terms—Group testing, probabilistic tests, sparsity recovery, compressed sensing, epidemiology.

I. INTRODUCTION

INVERSE problems, with the goal of recovering a signal from partial and noisy observations, come in many different formulations and arise in many applications. One important property of an inverse problem is to be well-posed, i.e., there should exist a unique and stable solution to the problem [2]. In this regard, prior information about the solution, like sparsity, can be used as a "regularizer" to transform an ill-posed problem to a well-posed one. In this work, we look at a particular inverse problem with less measurements than the number of unknowns (ill-posed) but with sparsity constraints on the solution. As will be explained in detail, the interesting aspect of this problem is that the sampling procedure is probabilistic and not fully known at recovery time.

Suppose that in a large set of items of size N, at most $K \ll N$ of them are defective and we wish to identify this small set of defective items. By testing each member of the set separately, we can expect the cost of the testing procedure to be large. If we could instead pool a number of items together and test the pool collectively, the number of tests required might be reduced. The ultimate goal is to construct a pooling design to identify the defective items while minimizing the number of tests. This is the main conceptual idea behind the classical *group testing* problem which was introduced by Dorfman [3] and later found applications in a variety of areas. The first important application of the idea dates back to World War II when it was suggested for syphilis screening. A few other examples of group testing applications include testing for defective items (e.g., defective light bulbs or resistors) as a part of industrial quality assurance [4], DNA sequencing [5], DNA library screening in molecular biology (see, e.g., [6]–[10] and the references therein), multi-access communication [11], data compression [12], pattern matching [13], streaming algorithms [14], software testing [15] and compressed sensing [16]. See the books by Du and Hwang for a detailed account of the major developments in this area [17], [18].

In a classical group testing setup, it is assumed that the reconstruction algorithm has full access to the sampling procedure, i.e., it knows which items participate in each pool. Moreover, if the tests are reliable, the existence of a defective item in a pool results in a positive test outcome. In an unreliable setting, the test results can be contaminated by false positives and/or false negatives. Compared to the reliable setting, special care should be taken to tackle this uncertainty in order to successfully identify the defective items. However, in some cases of interest, there can exist other types of uncertainty that challenge the recovery procedure.

In this work, we investigate the group testing problem with probabilistic tests. In this setting, a defective item which participate in a pool can be *inactive*, i.e., the test result of a pool can be negative despite containing some defective items. Therefore, a negative test result does not indicate that all

M. Cheraghchi is with the Department of Computer Science, University of Texas at Austin, USA (email: mahdi@cs.utexas.edu).

A. Hormati, A. Karbasi, and M. Vetterli are with the School of Computer and Communication Sciences, EPFL, Switzerland (emails: {ali.hormati, amin.karbasi, martin.vetterli}@epfl.ch). M. Vetterli is also with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, USA.
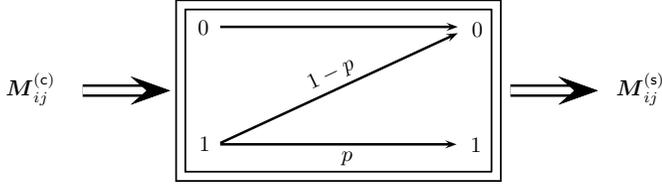
Fig. 1. Each element of the sampling matrix $M_{ij}^{(\mathrm{s})}$ is generated independently from the corresponding element of the contact matrix $M_{ij}^{(\mathrm{c})}$ by passing through a channel. The zeros in the contact matrix remain zeros in the sampling matrix while the ones are converted to zeros with probability $1-p$.

Fig. 2. Collective sampling using agents in viral epidemics. ($\bullet$) symbols represent healthy people while ($\otimes$) symbols indicate infected ones. The dashed lines connect the individuals contacted by the agents. An agent may remain healthy despite having contact with some infected people.

the items in the corresponding pool are non-defective with certainty. We follow a probabilistic approach to model the activity of the defective items, i.e., each defective item is active independently in each pool with probability p. Therefore, the tests contain uncertainty not in the sense of false positives or false negatives, but in the sense of the underlying probabilistic testing procedure. More precisely, let us denote by $M^{(\mathrm{c})}$ the designed contact matrix which indicates the items involved in each pool, i.e.,

$$M_{ij}^{(\mathrm{c})} = \begin{cases} 1 & \text{if test } i \text{ includes item } j \\ 0 & \text{otherwise.} \end{cases}$$

The probabilistic tests are then given by

$$y = M^{(\mathrm{s})} \cdot x$$

where $M^{(\mathrm{s})}$ denotes the probabilistic sampling matrix, x is the sparse input vector and y denotes the vector of test results. Each element of the contact matrix $M^{(\mathrm{c})}$ is independently mapped to the corresponding element of the sampling matrix $M^{(\mathrm{s})}$ by passing through the channel shown in Figure 1 [19]. In fact, the zeros in the contact matrix remain zeros in the sampling matrix while the ones are mapped to zeros with probability $1-p$.

In this work, our goal is to design efficient sampling and recovery mechanisms to successfully identify the sparse vector x, despite the partially unknown testing procedure given by the sampling matrix $M^{(\mathrm{s})}$. Our interest is in non-adaptive sampling procedures in which the sampling strategy (i.e., the contact matrix) is designed before seeing the test outcomes. In our analysis, we consider two different design strategies: In the *per-instance* design, the sampling procedure should be suitable for a fixed set of defective items with overwhelming probability while in the *universal* design, it should be appropriate for all possible sets of defective items. We show that $M = O\left(K \log(N)/p^3\right)$ tests are sufficient for successful recovery in the per-instance scenario while we need $M = O\left(K^2 \log(N/K)/p^3\right)$ tests for the universal design. Moreover, the defective items can be recovered by a simple recovery algorithm with complexity of $O(MN)$. For a constant parameter p, the bounds on the number of measurements are asymptotically tight up to logarithmic factors. This is simply because standard group testing that corresponds
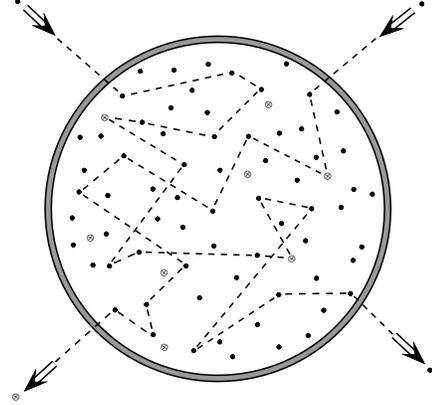
to the case $p = 1$ requires $M = \Omega(K^2 \log_K(N))$ non-adaptive measurements in the universal setting (cf. [17, Ch 7]) and $M = \Omega(K \log(N/K))$ measurements in the per-instance scenario (by a "counting argument").

The above-mentioned probabilistic sampling procedure can well model the sampling process in an epidemiology application, where the goal is to successfully identify a sparse set of virally-infected people in a large population with a few collective tests. In viral epidemics, one way to acquire collective samples is by sending "agents" inside the population whose task is to contact people. Once an agent makes contact with an "infected" person, there is a *chance* that he gets infected, too. By the end of the testing procedure, all agents are gathered and tested for the disease. Note that, when an agent contacts an infected person, he will get infected with a certain probability typically smaller than one. Hence, it may well happen that an agent's result is negative (meaning that he is not infected) despite a contact with some infected person. One can assume that each agent has a *log file* by which one can figure out with whom he has made contact. One way to implement the log in practice is to use identifiable devices (for instance, cell phones) that can exchange unique identifiers when in range. This way, one can for instance ask an agent to randomly meet a certain number of people in the population and at the end, learn which individuals have been met from the data gathered by the device. However, one should assume that when an agent gets infected, the resulting infection will not be contagious, i.e., an agent never infects other people[1]. In the above model, the agents can in fact take many forms, including people who happen to be in contact with random individuals within the population (e.g., cashiers, bus drivers, etc.).

The model explained using the epidemiology example above can in fact capture a broader range of settings, and in particular, any group testing problem where items can be defective

---

[1]This assumption is reasonable with certain diseases when there is an incubation time.

with a certain probability. An example of such a setting is testing for faulty components (or modules) in digital logic systems (e.g., an integrated circuit). This can be modeled through a probabilistic setting where the probability p denotes the percentage of time that a faulty component does not work correctly. In this way, one can use our group testing results to efficiently localize the few unreliable circuitry elements out of a large set of components on a given chip. One should note that this model generalizes the classical application of group testing for fault detection in electronic circuits where components are assumed to be either fully reliable or fully unreliable (see [20]).

The organization of this paper is as follows. We first give an overview of the related work in Section II which is then followed in Section III by a more precise formulation of our problem. In order to solve the original stochastic problem, we first solve an adversarial variation of it in Section IV which we find more convenient to work with. Then, in Section V and by using the results obtained from the adversarial setting, we design sensing and recovery procedures to efficiently solve the original stochastic problem. In Section VI, we provide a systematic design procedure which provides us with the exact value for the number of tests, along with the other necessary parameters, as a function of the desired probability of unsuccessful reconstruction. We evaluate the design by doing a set of numerical experiments. The paper is summarized in Section VII.

II. RELATED WORK

A large body of work in classical group testing has focused on combinatorial design, *i.e.*, construction of contact matrices that satisfy a disjunctness property (the exact definition will be provided in Section IV). Matrices that have this property are of significant interest since they ensure identifiability of defective items and moreover, they lead to efficient recovery algorithms. This property has been extensively studied in [21]–[25]. By using probabilistic methods, authors in [21] developed upper and lower bounds on the number of tests/rows for the contact matrix to be K-disjunct. More precisely, they showed that the number of rows should at least scale asymptotically as $O(K^2 \log N / \log K)$ for exact reconstruction with worst case input. On the other hand, a randomly generated matrix will be K-disjunct with high probability if the number of rows scales as $O(K^2 \log(N/K))$ [22]. Having a K-disjunct matrix, one can devise an efficient reconstruction algorithm to identify up to K defective items. This is true if the reconstruction algorithm fully knows the sampling procedure. However, in our scenario, the decoder has to cope simultaneously with two sources of uncertainty, the unknown group of defective items and the partially unknown (or stochastic) sampling procedure. For this reason we need to use a more general form of disjunctness property.

We should also point out the relationship between our setup and the compressed sensing (CS) framework [26], [27]. In CS, a random projection of a sparse signal is given and the goal is to find the position as well as the value of the non-zero entries of the input signal while keeping the number of projections to

a minimum. Exploiting the similarity between group testing and CS, new recovery algorithms for sparse signal recovery have been proposed in [28] and [29]. Although in CS the goal is to measure and reconstruct sparse signals with few measurements, it differs in significant ways from our setup. In CS, it is typically assumed that the decoder knows the measurement matrix a priori[2]. However, this is not the case in our setup. In other words, by using the language of compressed sensing, the measurement matrix might be "noisy" and not precisely known to the decoder. As it turns out, by using a sufficient number of tests this issue can be resolved. Another difference is that in CS, the input and the measurement matrix are real valued and operations are performed on real numbers whereas in our case, the input vector, the measurement matrix and the arithmetic are all boolean.

Recently, the authors in [19] investigated the probabilistic testing model that we consider in this paper from an information theoretic perspective. Unlike our combinatorial approach, they use information theoretic techniques to obtain bounds on the required number of tests. Namely, they get $M = O(K^2 (\log N)/p^2)$ and $M = O(K(\log N)/p^2)$ measurements for universal and per-instance scenarios, respectively, which is asymptotically comparable to what we obtain in this work (for a fixed p and $K \ll N$). To achieve the bounds, they consider typical set decoding as the reconstruction method. However, they do not provide a practical, low complexity decoding algorithm for the reconstruction.

Another work that is relevant to ours is [35] that considers group testing under a "dilution" effect. This model is targeted for biological experiments where combining items in a group may cause defected items go undetected when the size of the group is large. In particular, their model assumes that each item is independently defected with a certain (fixed) probability, and a defected item in a group of size t affects the test with a probability proportional to $1/t$ (thus, a "diluted" group with few defectives becomes more likely to test negative as its size grows). They analyze the number of required tests using a simple (but sub-optimal) test design originally proposed by Dorfman [3].

III. PROBLEM DEFINITION

To model the problem, we enumerate the population from 1 to N and the tests from 1 to M. Let the nonzero entries of $\boldsymbol{x} := (x_1, x_2, \ldots, x_N) \in \mathbb{F}_2^N$ indicate the defective items within the population, where $\mathbb{F}_2$ is the finite field of order 2. Moreover, we assume that $\boldsymbol{x}$ is a K-sparse vector, i.e., it has at most K entries equal to one (corresponding to the defective items). We refer to the *support set* of $\boldsymbol{x}$, denoted by $\mathrm{supp}(\boldsymbol{x})$, as the set which contains positions of the nonzero entries.

As is typical in the literature of group testing, we introduce an $M \times N$ boolean *contact* matrix $\boldsymbol{M}^{(\mathrm{c})}$ to model the set of non-adaptive tests. We set $\boldsymbol{M}_{ij}^{(\mathrm{c})}$ to one if and only if the ith test contains the jth item. The matrix $\boldsymbol{M}^{(\mathrm{c})}$ only shows which

---

[2]There are, however, works that consider compressed sensing under small perturbations of the measurement matrix (cf. [30]). A large body of the compressed sensing literature considers a noise model where the measurement outcomes are perturbed by a real-valued noise vector, while the measurement matrix is exact. See, for example, [31]–[34] and the references therein.

tests contain which items. In particular, it does not indicate whether the tests eventually get affected by the defective items. Let us assume that when a test contains a set of defective items, each of them makes the test positive independently with probability p, which is a fixed parameter that we call the *activation probability*. Therefore, the real *sampling* matrix $M^{(s)}$ can be thought of as a variation of $M^{(c)}$ in the following way: Each nonzero entry of $M^{(c)}$ is flipped to 0 independently with probability $1-p$. Then, the resulting matrix $M^{(s)}$ is used just as in classical group testing to produce the test vector $y \in \mathbb{F}_2^M$

$$y = M^{(s)} \cdot x$$

where the arithmetic is boolean, i.e., multiplication with the logical AND and addition with the logical OR.

The contact matrix $M^{(c)}$, the test vector y, the upperbound on the number of nonzero entries K, and the activation probability p are known to the decoder, whereas the sampling matrix $M^{(s)}$ (under which the collective samples are taken) and the input vector x are unknown. The task of the decoder is to identify the K nonzero entries of x based on the known parameters.

Example. As a toy example, consider a population with 6 items where only two of them (items 3 and 4) are defective. We do a set of three tests, where the first one contains items $1, 3, 5$, the second one contains items $2, 4, 6$, and the third one contains items $2, 3, 5, 6$. Therefore, the contact matrix and the input vector have the following form

$$x = (\ 0 \ \ 0 \ \ 1 \ \ 1 \ \ 0 \ \ 0 \ )^T,$$
$$M^{(c)} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Let us assume that only the second test result is positive. This means that the test vector is

$$y = (\ 0 \ \ 1 \ \ 0 \ )^T.$$

As we can observe, there are many possibilities for the sampling matrix, all of the following form:

$$M^{(s)} = \begin{pmatrix} ? & 0 & ? & 0 & ? & 0 \\ 0 & ? & 0 & ? & 0 & ? \\ 0 & ? & ? & 0 & ? & ? \end{pmatrix}$$

where the question marks are 0 with probability $1-p$ and 1 with probability p. It is the decoder's task to figure out which combinations make sense based on the outcome vector. For example, the following matrices and input vectors fit perfectly

with y:

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

More formally, the goal of our scenario is two-fold:
1) Designing the contact matrix $M^{(c)}$ so that it allows unique reconstruction of sparse input x from outcome y with overwhelming probability $1 - o(1)$ over the randomness of the sampling matrix $M^{(s)}$.
2) Proposing a recovery algorithm with low computational complexity.

We present a probabilistic approach for designing contact matrices suitable for our problem setting, along with a simple decoding algorithm for reconstruction. Our approach is to first introduce a rather different setting for the problem that involves no randomness in the way the defective items become active. Namely, in the new setting an adversary can arbitrarily decide whether a certain contact with a defective item results in a positive test result or not, and the only restriction on the adversary is on the total amount of inactive contacts being made. The reason for introducing the adversarial problem is its combinatorial nature that allows us to use standard tools and techniques already developed in combinatorial group testing. Fortunately, it turns out that by choosing a carefully-designed value for the total amount of inactive contacts based on the parameters of the system, solving the adversarial variation is sufficient for the original (stochastic) problem.

Our task is then to design contact matrices suitable for the adversarial problem. We give a probabilistic construction of the contact matrix in Section V. The probabilistic construction requires each test to independently contact the items with a certain well-chosen probability. This construction ensures that the resulting data gathered at the end of the experiment can be used for correct identification of the defective items with overwhelming probability, provided that the number of tests is sufficiently large. In our analysis, we consider two different design strategies
- **Per-Instance Design**: The contact matrix is suitable for every arbitrary, but a priori fixed, sparse input vector with overwhelming probability.
- **Universal Design**: The contact matrix is suitable for all sparse input vectors with overwhelming probability.

Based on the above definitions, the contact matrix constructed for the per-instance scenario, once fixed, may fail to distinguish between all pairs of sparse input vectors. On the other hand, in the universal design, one can use a single contact

matrix to successfully measure all sparse input vectors with a very high probability of success. Our results show that $M = O\left(K \log(N)/p^3\right)$ tests are sufficient for successful recovery in the per-instance scenario while we need $M = O\left(K^2 \log(N/K)/p^3\right)$ tests for the universal design.

Remark: As is customary in the standard group testing literature, we think of the sparsity K as a parameter that is noticeably smaller than the population size N, for example, one may take $K = O(N^{1/4})$. Indeed, if K becomes comparable to N, there would be little point in using a group testing scheme and in practice, for large K it is generally more favorable to perform trivial tests on the items.

IV. ADVERSARIAL SETTING

The problem described in Section III has a stochastic nature, i.e., the sampling matrix is obtained from the contact matrix through a random process. In this section, we introduce an adversarial variation of the problem whose solution leads us to the solution for the original stochastic problem.

In the adversarial setting, the sampling matrix is obtained from the contact matrix by flipping up to e arbitrary entries to 0 on the support (i.e., the set of nonzero entries) of each column of $M^{(c)}$. The goal is to be able to exactly identify the sparse input vector despite the perturbation of the contact matrix and regardless of the choice of the flipped entries. Note that the classical group testing problem corresponds to the special case $e = 0$. Thus, the only difference between the adversarial problem and the stochastic one is that in the former, the flipped entries of the contact matrix are chosen arbitrarily (as long as there are not too many flips) while in the latter, they are chosen according to a specific random process.

It turns out that the combinatorial tool required for solving the adversarial problem is closely related to the notion of *disjunct* matrices that is well studied in the group testing literature [17]. The formal definition is as follows.

Definition 1: A boolean matrix M with N columns $M_1, \ldots, M_N$ is called (K, e)-disjunct if, for every subset S of the columns with $|S| \leq K$, and every $i \in [N]$, we have

$$\left| \mathsf{supp}(M_i) \setminus \left(\bigcup_{j \in S \setminus \{i\}} \mathsf{supp}(M_j) \right) \right| > e$$

where $\mathsf{supp}(M_i)$ denotes the support set of the column M_i and $\setminus$ is the set difference operator. In words, this operation counts the number of nonzero positions in the column M_i for which all columns with index in the set S have zeros.

Note that the special case of $(K, 0)$-disjunct matrices corresponds to the classical notion of K-disjunct matrices which is essentially equivalent to strongly selective families and superimposed codes (see [36]). Moreover, when all columns of the matrix have the same Hamming weight t, a $(K+1, 2t/3)$-disjunct matrix turns out to be equivalent to t-majority K-strongly selective families that are defined in [37] (where each row of the matrix defines the characteristic vector of a set in the family). This notion is known to be useful for construction of non-adaptive compressed sensing schemes [37].

The following proposition shows the relationship between contact matrices suitable for the adversarial problem and disjunct matrices.

Proposition 2: Let M be a (K, e)-disjunct matrix. Then taking M as the contact matrix solves the adversarial problem for K-sparse vectors with error parameter e. Conversely, any matrix that solves the adversarial problem must be $(K-1, e)$-disjunct.

Proof: Let M be a (K, e)-disjunct matrix and consider K-sparse vectors x and x' supported on different subsets S and S', respectively. Take an element $i \in S'$ which is not in S. By Definition 1, we know that the column M_i has more than e entries on its support that are not present in the support of any M_j, $j \in S$. Therefore, even after e bit flips in M_i, at least one entry in its support remains that is not present in the test outcome of x', and this makes x and x' distinguishable.

For the reverse direction, suppose that M is not $(K-1, e)$-disjunct and take any i and a subset S with $|S| \leq K-1$, $i \notin S$ which demonstrates a counterexample for M being $(K-1, e)$-disjunct. Consider K-sparse vectors x and x' supported on S and $S \cup \{i\}$, respectively. An adversary can flip up to e bits on the support of M_i from 1 to 0, leave the rest of M unchanged, and ensure that the test outcomes for x and x' coincide. Thus M is not suitable for the adversarial problem. ∎

A. Distance Decoder

Proposition 2 shows that a (K, e)-disjunct contact matrix can *combinatorially* distinguish between K-sparse vectors in the adversarial setting with error parameter e. In the following, we show that there exists a much simpler decoder for this purpose.

Distance decoder: For any column c_i of the contact matrix $M^{(c)}$, the decoder verifies the following:

$$|\mathsf{supp}(c_i) \setminus \mathsf{supp}(y)| \leq e \tag{1}$$

where y is the vector consisting of the test outcomes. The coordinate x_i is decided to be nonzero if and only if the inequality holds.

The above decoder is a straightforward generalization of a standard decoder that is used in classical group testing. The standard decoder chooses those columns of the measurement matrix whose supports are fully contained in the measurement outcomes (see [17, Ch 7]).

Proposition 3: Let the contact matrix $M^{(c)}$ be (K, e)-disjunct. Then, the distance decoder correctly identifies the correct support of any K-sparse vector in the adversarial setting with error parameter e.

Proof: Let x be a K-sparse vector and $S := \mathsf{supp}(x)$, $|S| \leq K$, and $M_S^{(c)}$ denote the corresponding set of columns in the sampling matrix. Obviously, all the columns in $M_S^{(c)}$ satisfy (1) (as no column is perturbed in more than e positions) and thus the reconstruction includes the support of x (this is true regardless of the disjunctness property of $M^{(c)}$). Now let the vector $\hat{y}$ be the bitwise OR of the columns in $M_S^{(c)}$ and assume that there is a column c of $M^{(c)}$ outside S that

satisfies (1). Thus, since

$$\mathsf{supp}(\boldsymbol{y}) \subseteq \mathsf{supp}(\hat{\boldsymbol{y}}),$$

we will have

$$|\mathsf{supp}(\boldsymbol{c}) \setminus \mathsf{supp}(\hat{\boldsymbol{y}})| \le e$$

which violates the assumption that $\boldsymbol{M}^{(\mathsf{c})}$ is (K, e)-disjunct for the support set S and the column $\boldsymbol{c}$ outside this set. Therefore, the distance decoder outputs the exact support of $\boldsymbol{x}$. ∎

Of course, posing the adversarial problem is interesting if it helps in solving the original stochastic problem from which it originates. In the next section, we show that this is indeed the case; and in fact the task of solving the stochastic problem reduces to that of the adversarial problem.

V. PROBABILISTIC DESIGN

In this section, we consider a probabilistic construction for $\boldsymbol{M}^{(\mathsf{c})}$, where each entry of $\boldsymbol{M}^{(\mathsf{c})}$ is set to 1 independently with probability $q = \alpha/K$, for a parameter α to be determined later, and 0 with probability $1 - q$. We will use standard arguments to show that, if the number of tests M is sufficiently large, then the resulting matrix $\boldsymbol{M}^{(\mathsf{c})}$ is suitable with all but a vanishing probability.

By looking carefully at the proof of Proposition 3, we see that there are two events that *may prevent* the distance decoder with error parameter e to successfully recover the input vector $\boldsymbol{x}$ with support on S:

1) There are more than e flips on the columns of the contact matrix in S.
2) There exists a column outside S where the (K, e)-disjunct property is violated.

Based on these observations, the number of tests required for building suitable contact matrices are given by the following theorem.

Theorem 4: Consider $M \times N$ contact matrices $\boldsymbol{M}^{(\mathsf{c})}$ constructed by the probabilistic design procedure. If $M = O\left(K \log(N)/p^3\right)$ for the per-instance scenario or $M = O\left(K^2 \log(N/K)/p^3\right)$ for the universal scenario, then the probability of failure for the reconstruction with the distance decoder goes to zero as $N \to \infty$.

Proof: Let e be the decision-making parameter of the distance decoder. We first find an upperbound for the number of bit flips in any column of the contact matrix $\boldsymbol{M}^{(\mathsf{c})}$. To this end, take any column $\boldsymbol{M}_i^{(\mathsf{c})}$ of $\boldsymbol{M}^{(\mathsf{c})}$. Each entry of the column $\boldsymbol{M}_i^{(\mathsf{c})}$ is flipped independently with probability $(1-p)q$ which, on average, results in $(1-p)qM$ bit flips per column. Let $e = (1 + \delta)(1 - p)qM$ for a constant $\delta > 0$. By Chernoff bounds (cf. [38]), the probability that the amount of bit flips exceeds e is at most

$$\exp\left(-\delta^2 (1-p)qM/(2+\delta)\right). \tag{2}$$

Second, we check the disjunctness property of the contact matrix for this parameter e. To this end, consider any set S of K columns of $\boldsymbol{M}^{(\mathsf{c})}$, and any column outside these, say the ith column where $i \notin S$. First we upper bound the probability of *failure* for this choice of S and i. That is, the probability that

the number of rows that have a 1 at the ith column and all-zeros at the positions corresponding to S is at most e. Clearly if this event happens the (K, e)-disjunct property is violated.

A row is *good* if at that row the ith column has a 1 but all the columns in S have zeros. For a particular row, the probability that the row is good is $q(1 - q)^K$ (using independence of the entries of the measurement matrix). Then failure corresponds to the event that the number of good rows is at most e. The distribution on the number of good rows is binomial with mean $\mu = q(1 - q)^K M$. Using the Chernoff bound and assuming that $e < \mu$ (we will choose α and δ to ensure this condition is satisfied), we have[3]

$$\text{failure probability} \le \exp\left(\frac{-(\mu - e)^2}{2\mu}\right)$$

$$= \exp\left(\frac{-Mq\left((1-q)^K - (1-p)(1+\delta)\right)^2}{2(1-q)^K}\right) \tag{3}$$

$$\overset{(a)}{\le} \exp\left(\frac{-Mq\left((1-2\alpha) - (1-p)(1+\delta)\right)^2}{2(1-\alpha/2)}\right)$$

$$\le \exp\left(-\frac{1}{2}Mq\left((1-2\alpha) - (1-p)(1+\delta)\right)^2\right)$$

$$= \exp(-M\gamma/K), \tag{4}$$

where we have defined

$$\gamma := \frac{\alpha}{2}\left((1-2\alpha) - (1-p)(1+\delta)\right)^2$$

The inequality (a) is due to the fact that $(1 - q)^K = (1 - \alpha/K)^K$ is always between $3^{-\alpha}$ and $2^{-\alpha}$, and in particular for $\alpha \in [0, 1]$, this range is strictly contained in $[1 - 2\alpha, 1 - \alpha/2]$. Note that by choosing the parameters α and δ sufficiently small, the quantity

$$\left((1-2\alpha) - (1-p)(1+\delta)\right)^2$$

in the exponent can be made arbitrarily close to p^2. As a concrete choice, however, we take $\delta := p/2$ and $\alpha := p/8$ which gives

$$\gamma = \frac{p}{16}\left(\frac{p}{4} + \frac{p^3}{4}\right)^2 = \Omega(p^3)$$

and, therefore,

$$\text{failure probability} \le 2^{-\Omega(Mp^3/K)}. \tag{5}$$

In order to calculate the number of tests, we consider per-instance and universal scenarios separately.

- *Per-Instance Scenario*

  For the per-instance scenario, the disjunctness property needs to hold only for a fixed set S, corresponding to the support of the fixed sparse vector that defines the instance. Therefore, we only need to apply the union bound over all possible choices of i for a fixed set S. From (5), the probability of coming up with a bad choice of $\boldsymbol{M}^{(\mathsf{c})}$ would thus be at most

$$N2^{-\Omega(Mp^3/K)}.$$

[3]The failure probability is at least 0.5 if $e \ge \mu$.

This probability vanishes for an appropriate choice of

$$M = \Theta\left(\frac{K \log N}{p^3}\right). \qquad (6)$$

At the same time, using (2) and the union bound, the probability that the amount of bit flips in any of the K columns in S exceeds e is upper bounded by

$$K \exp\left(-\delta^2(1-p)qM/(2+\delta)\right)$$
$$= K \exp\left(-\Omega\left(\frac{\alpha\delta^2(1-p)\log(N)}{(2+\delta)p^3}\right)\right),$$

which is vanishing (i.e., $o(1)$) assuming the constant behind the $\Theta(\cdot)$ notion in (6) is sufficiently large. Therefore, the distance decoder successfully decodes the input vector $\boldsymbol{x}$ with probability $1 - o(1)$ in the per-instance scenario with $M = O\left(K \log(N)/p^3\right)$ tests.

- *Universal Scenario*

  In this case, we apply the union bound over all possible choices of S and i. Using (4), the probability of coming up with a bad choice of $\boldsymbol{M}^{(\text{c})}$ is at most $N\binom{N}{K}\exp\left(-M\gamma/K\right)$. This probability vanishes for an appropriate choice of

  $$M = \Theta\left(\frac{K^2 \log(N/K)}{\gamma}\right) = \Theta\left(\frac{K^2 \log(N/K)}{p^3}\right).$$

  At the same time, using (2) and the union bound, the probability that the amount of bit flips in any of the N columns of the contact matrix exceeds e is upper bounded by

  $$N \exp\left(-\delta^2(1-p)qM/(2+\delta)\right)$$
  $$= N \exp\left(-\frac{\alpha\delta^2(1-p)K\log(N/K)}{(2+\delta)\gamma}\right) = o(1).$$

  Therefore, with $M = O\left(K^2 \log(N/K)/p^3\right)$ tests, the probabilistic design constructs a contact matrix such that the distance decoder is able to decode all sparse input vectors $\boldsymbol{x}$ with probability $1 - o(1)$. ∎

The probabilistic construction results in a rather sparse contact matrix, namely, one with density $O(1/K)$ that decays with the sparsity parameter K. In the following, we show that sparsity is necessary for the probabilistic construction to work.

Proposition 5: Let $\boldsymbol{M}$ be an $M \times N$ boolean random matrix, where $M = O\left(K^2 \log(N/K)\right)$ or $M = O\left(K \log(N)\right)$ for an integer $K > 0$, which is constructed by setting each entry independently to 1 with probability q. Then either $q = O\left(\log K/K\right)$ or otherwise the probability that $\boldsymbol{M}$ is (K, e)-disjunct (for any $e \geq 0$) approaches to zero as N grows.

Proof: Suppose that $\boldsymbol{M}$ is an $M \times N$ matrix that is (K, e)-disjunct. Observe that, for any integer $t \in (0, K)$, if we remove any t columns of $\boldsymbol{M}$ and all the rows on the support of those columns, the matrix must remain $(K - t, e)$-disjunct. This is because any counterexample for the modified matrix being $(K - t, e)$-disjunct can be extended to a counterexample for $\boldsymbol{M}$ being (K, e)-disjunct by adding back the removed columns and rows.

Now consider any t columns of $\boldsymbol{M}$, and denote by M_0 the number of rows of $\boldsymbol{M}$ at which the entries corresponding to the chosen columns are all zeros. The expected value of M_0 is $(1 - q)^t M$. Moreover, for every $\delta \in (0, 1)$ we have

$$\Pr\left[M_0 > (1 + \delta)(1 - q)^t M\right] \leq \exp\left(-\frac{\delta^2}{3}(1 - q)^t M\right)$$
$$(7)$$

by the Chernoff bound. Let t_0 be the largest integer for which

$$(1 + \delta)(1 - q)^{t_0} M \geq \log N.$$

If $t_0 < K - 1$, we let $t := t_0 + 1$ above, and this makes the right hand side of (7) upper bounded by $o(1)$. So with probability $1 - o(1)$, the chosen t columns of $\boldsymbol{M}$ will keep M_0 at most $(1 + \delta)(1 - q)^t M$. Removing those columns and all the rows on the support of these columns leaves the matrix $(K - t, e)$-disjunct, which obviously requires at least $\log N$ rows as even a $(1, 0)$-disjunct matrix needs so many rows. Therefore, we must have

$$(1 + \delta)(1 - q)^t M \geq \log N.$$

However, this inequality is not satisfied by the assumption on t_0. So if $t_0 < K - 1$, little chance remains for $\boldsymbol{M}$ to be (K, e)-disjunct for any $e \geq 0$. Therefore, we should have $t_0 \geq K - 1$. Using the condition on t_0, we have

$$(1 + \delta)(1 - q)^{K-1} M \geq \log N.$$

This is equivalent to

$$q \leq \frac{\log\left(M(1 + \delta)/\log N\right)}{K - 1}$$

which for $M = O\left(K^2 \log(N/K)\right)$ or $M = O\left(K \log(N)\right)$ gives $q = O\left(\log K/K\right)$. ∎

In summary, our results indicate that for both the per-instance and universal settings, the activation probability p increases the upper bound on the number of tests by a factor of $1/p^3$. Moreover, we can use the simple distance decoder to recover the unknown input vector with the complexity of $O(MN)$. However, in order for the probabilistic design to work, we should choose a flip probability q such that $q = O\left(\log K/K\right)$. In fact, our choice of $q = \alpha/K$ for a constant α satisfies this requirement.

VI. SYSTEM DESIGN AND SIMULATION RESULTS

In this section, we provide a systematic design procedure which gives us the number of tests necessary for the decoding process to be successful. While the design procedure applies to both per-instance and universal scenarios, the numerical simulation result is provided only for the per-instance setting, since evaluating the universal design requires to test all possible inputs which is computationally prohibitive.

According to the discussion in Section V, there are two types of failure events which we want to avoid in designing the contact matrix $\boldsymbol{M}^{(\text{c})}$. The first failure event, denoted as f_1, happens when the number of bit flips in a column is not tolerable by the contact matrix and the second event, denoted as f_2, relates to the violation of the disjunctness property of the matrix. The inputs to the design procedure are N, K,

p, p_{f_1} and p_{f_2}, where the last two parameters denote the maximum tolerable probability for the first and second failure events, respectively. Then, the design procedure should provide us with the quantities M, q and e, which are the required parameters to setup the sensing and recovery algorithms.

Let us summarize the results of the probabilistic design of Section V. First we define η from (3) as

$$\eta = q \frac{\left((1-q)^K - (1-p)(1+\delta)\right)^2}{2(1-q)^K}. \tag{8}$$

Then, for the per-instance scenario (which is denoted by (i)), we have

$$p_{f_1}^{(i)} \leq 1 - \left[1 - \exp\left(-(1-p)\, q\, M^{(i)}(\log(1+\delta)^{(1+\delta)} - \delta)\right)\right]^K,$$

$$p_{f_2}^{(i)} \leq N \exp\left(-M^{(i)}\eta\right),$$
$$e^{(i)} = (1+\delta)(1-p)\, q M^{(i)}.$$

For the universal strategy (which is denoted by (u)), we have

$$p_{f_1}^{(u)} \leq 1 - \left[1 - \exp\left(-(1-p)\, q\, M^{(u)}(\log(1+\delta)^{(1+\delta)} - \delta)\right)\right]^N,$$

$$p_{f_2}^{(u)} \leq N \binom{N}{K} \exp\left(-M^{(u)}\eta\right),$$
$$e^{(u)} = (1+\delta)(1-p)\, q M^{(u)}.$$

Note that since the first failure event happens independently on the columns, we have used a more precise expression for this failure probability which does not use the union bound, and also makes use of the exact expression for the Chernoff bound.

Let us provide the details of the design for the per-instance scenario; The universal design follows the same lines. For any fixed value of α, $\delta > 0$ is the only parameter which should be determined such that the failure probabilities fall below the maximum tolerable values. To this end, we initialize the value of δ to zero and increase it in small steps up to the value $\delta_{\max}$. Given that $e = (1-p)(1+\delta)\, q M^{(i)}$ and $\mu = q(1-q)^K M^{(i)}$ and under the condition that $e < \mu$, we have

$$\delta_{\max} = (1-q)^K/(1-p) - 1.$$

For any value of $\delta < \delta_{\max}$ and given the maximum tolerable probability for the second failure event p_{f_2}, the number of tests are computed as

$$M^{(i)} = \eta^{-1} \log\left(N/p_{f_2}\right)$$

where η is defined in (8). This is then used to compute the corresponding probability for the first failure event as

$$p_{f_1}^{(i)} = 1 - \left[1 - \exp\left(-(1-p)\, q\, M^{(i)}(\log(1+\delta)^{(1+\delta)} - \delta)\right)\right]^K.$$
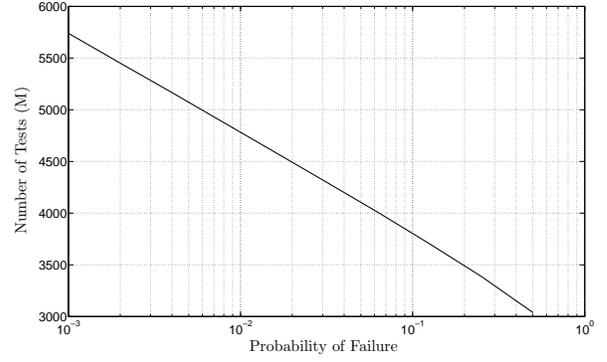

Fig. 5. The number of tests as a function of the failure probability for the per-instance scenario. The parameters are $N = 100'000$, $K = 10$ and activation probability $p = 0.8$.

We continue increasing δ until $p_{f_1}^{(i)}$ falls below the maximum tolerable probability for the first failure event p_{f_1}. This provides us with the number of tests $M^{(i)}$ and the error parameter $e^{(i)}$ for the chosen value of α. This whole process is continued for different values of α in the range $[0, \alpha_{\max}]$. At the end, we find the value of α which results in the minimum number of tests for the given value of p. This provides us with the number of tests and the decision parameter for the distance decoder.

Figures 3(a) and 3(b) show the number of tests, for universal and per-instance strategies, as a function of the parameter α and the activation probability p. The population size is $N = 100'000$, the number of defective items is $K = 10$ and the maximum tolerable probabilities for the two failure events are set to $p_{f_1} = p_{f_2} = 0.001$. Note that the number of tests for the per-instance scenario is much less than the universal scenario and moreover, it allows us to have designs appropriate for smaller activation probabilities. The black curve in each figure connects the points with minimum number of tests for each value of the activation probability p, which in turn provides us with the appropriate value for the parameter α. The black curves are extracted and shown separately in Figure 3(c). In Figures 4(a), 4(b) and 4(c), we show the output of the design procedure for $N = 100'000'000$, $K = 500$ and $p_{f_1} = p_{f_2} = 0.001$.

In Figure 5, we set $N = 100'000$, $K = 10$ and $p = 0.8$ and use the design procedure to plot the number of tests as a function of the probability of failure in the per-instance strategy. Then, in Figure 6, we run a numerical experiment with the same values for the parameters N, K and p to assess the performance of the recovery algorithm, with the results averaged over 4000 trials. We set the parameters e and α for the numerical experiment equal to those which give us the probability of failure of 0.5 in Figure 5 (which are $e = 40$ and $\alpha = 0.44$) and change the number of tests. Note that although we expect a failure probability of around 0.5 for $M = 3000$ tests according to Figure 5, the recovery performance is much better in numerical simulations. This can be explained by noting that the upper bounds for the failure probabilities are not tight in general.
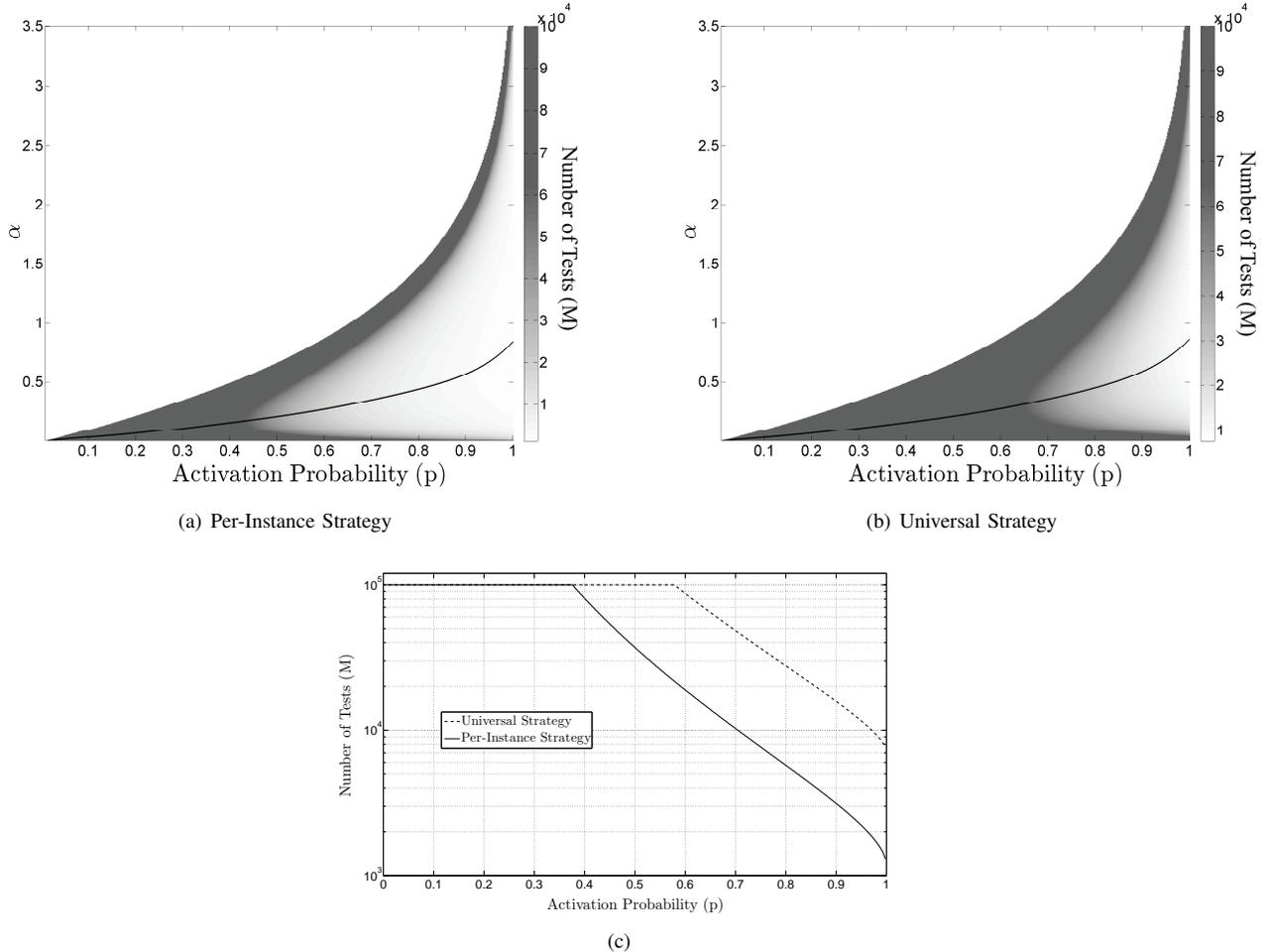
(a) Per-Instance Strategy

(b) Universal Strategy

(c)

Fig. 3. The number of tests (M) as the output of the design procedure for (a) per-instance and (b) universal strategies, as a function of the parameter α and the activation probability p. The parameters are set to $N = 100'000$, $K = 10$ and $p_{f_1} = p_{f_2} = 0.001$. The black curves provide us with the value of α which gives the minimum number of tests for each value of the activation probability p. (c) The minimum number of tests (corresponding to the black curves in (a) and (b)) for universal and per-instance strategies, as a function of the activation probability p.
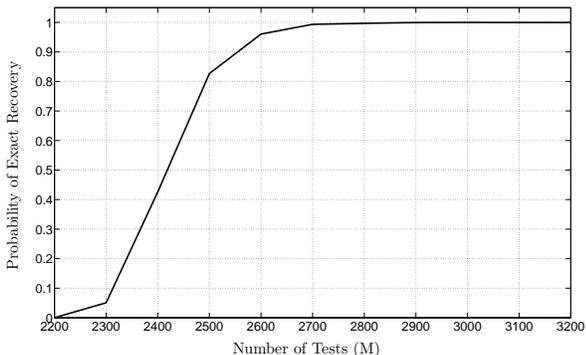

Fig. 6. The simulated probability of exact recovery for the per-instance strategy for $N = 100'000$, $K = 10$ and $p = 0.8$, averaged over 4000 trials. Although the design procedure expects a high probability of failure for $M = 3000$ tests (see Figure 5), the recovery performance is much better in numerical simulations.

VII. Conclusion

We studied the problem of identifying a small number of defective items among a large population, using collec-

tive samples. With the viral epidemic application in mind, we investigated the case where the recovery algorithm may possess only partial knowledge about the sampling process, in the sense that the defective items can become inactive in the test results. We showed that by using a probabilistic model for the sampling process, one can design a non-adaptive contact matrix which leads to the successful identification of the defective items with overwhelming probability $1 - o(1)$. We considered two strategies for the design procedure. In per-instance design, the contact matrix is suitable for each sparse input vector with overwhelming probability while in universal design, this is true for all sparse inputs. To this end, we proposed a probabilistic design procedure which requires a "small" number of tests to single out the sparse vector of defective items. More precisely, we showed that for an activation probability p, the number of tests sufficient for identification of up to K defective items in a population of size N is given by $M = O(K \log(N)/p^3)$ for the per-instance scenario and $M = O(K^2 \log(N/K)/p^3)$ for the universal scenario. Moreover, we proposed a simple decoder which is able to successfully identify the defective items with
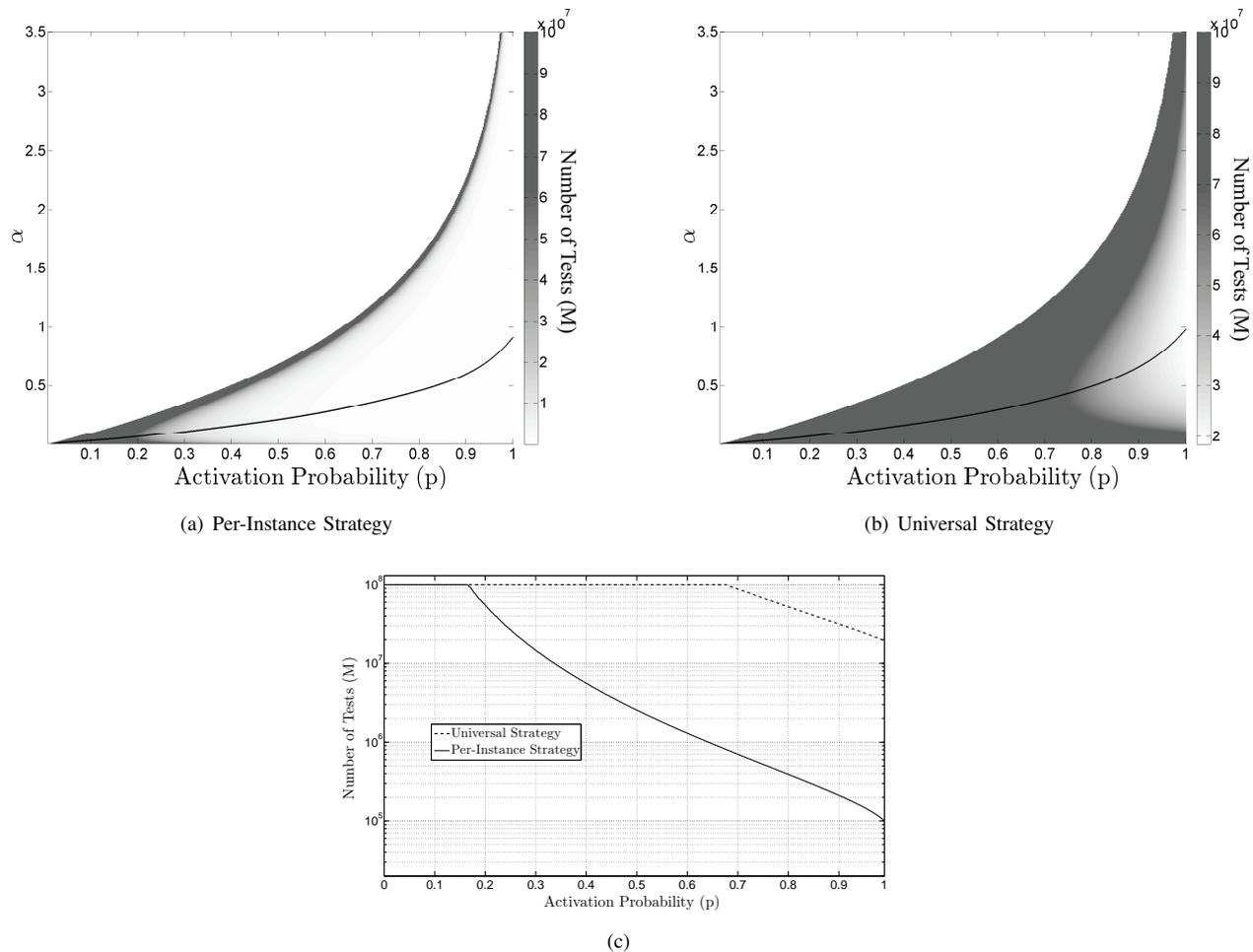
(a) Per-Instance Strategy

(b) Universal Strategy

(c)

Fig. 4. The number of tests (M) as the output of the design procedure for (a) per-instance and (b) universal strategies, as a function of the parameter α and the activation probability p. The parameters are set to $N = 100'000'000$, $K = 500$ and $p_{f_1} = p_{f_2} = 0.001$. The black curves provide us with the value of α which gives the minimum number of tests for each value of the activation probability p. (c) The minimum number of tests (corresponding to the black curves in (a) and (b)) for universal and per-instance strategies, as a function of the activation probability p.

complexity of $O(MN)$. Finally, we provided a systematic design procedure which gives the number of tests M, along with the design parameters α and e, required for successful recovery. As expected, the numerical experiments showed that the number of tests provided by the design procedure overestimates the true one required to achieve a specified probability of failure. As a complement to this work, one can also consider the effects of false positives and false negatives on the required number of tests. We leave this issue for future work.

REFERENCES

[1] M. Cheraghchi, A. Hormati, A. Karbasi, and M. Vetterli, "Compressed sensing with probabilistic measurements: A group testing solution," in *47th Annual Allerton Conference on Communication, Control, and Computing*, 2009.

[2] J. Hadamard, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*. Yale Univ. Press, New Haven, 1923.

[3] R. Dorfman, "The detection of defective members of large populations," *Annals of Mathematical Statistics*, vol. 14, pp. 436–440, 1943.

[4] M. Sobel and P. Groll, "Group-testing to eliminate efficiently all defectives in a binomial sample," *Bell Systems Technical Journal*, vol. 38, pp. 1179–1252, 1959.

[5] P. Pevzner and R. Lipshutz, "Towards DNA sequencing chips," in *19th International Symposium on Mathematical Foundations of Computer Science*, vol. 841, pp. 143–158, 1994.

[6] H. Ngo and D. Du, "A survey on combinatorial group testing algorithms with applications to DNA library screening," *DIMACS Series on Discrete Mathematical Problems and Theoretical Computer Science*, vol. 55, pp. 171–182, 2000.

[7] A. Schliep, D. Torney, and S. Rahmann, "Group testing with DNA chips: Generating designs and decoding experiments," in *IEEE Computer Society Bioinformatics Conference*, vol. 2, pp. 84–91, 2003.

[8] A. Macula, "Probabilistic nonadaptive group testing in the presence of errors and DNA library screening," *Annals of Combinatorics*, vol. 3, no. 1, pp. 61–69, 1999.

[9] A. Macula, "Probabilistic nonadaptive and two-stage group testing with relatively small pools and DNA library screening," *Journal of Combinatorial Optimization*, vol. 2, pp. 385–397, 1999.

[10] Y. Cheng and D. Z. Du, "New constructions of one- and two-stage pooling designs," *Journal of Computational Biology*, vol. 15, no. 2, pp. 195–205, 2008.

[11] J. Wolf, "Born-again group testing: multiaccess communications," *IEEE Transactions on Information Theory*, vol. 31, pp. 185–191, 1985.

[12] E. S. Hong and R. E. Ladner, "Group testing for image compression," *IEEE Transactions on Image Processing*, vol. 11, no. 8, pp. 901–911, 2002.

[13] R. Clifford, K. Efremenko, E. Porat, and A. Rothschild, "k-mismatch with don't cares," in *Proceedings of the 15th European Symposium on Algorithm (ESA)*, vol. 4698 of *Lecture Notes in Computer Science*, pp. 151–162, 2007.

[14] G. Cormode and S. Muthukrishnan, "What's hot and what's not: tracking most frequent items dynamically," *ACM Transactions on Database Systems*, vol. 30, no. 1, pp. 249–278, 2005.

[15] A. Blass and Y. Gurevich, "Pairwise testing," *Bulletin of the European Association for Theoretical Computer Science*, vol. 78, pp. 100–132, 2002.

[16] G. Cormode and S. Muthukrishnan, "Combinatorial algorithms for compressed sensing," in *Information Sciences and Systems*, pp. 198–201, Mar. 2006.

[17] D. Z. Du and F. Hwang, *Combinatorial Group Testing and its Applications*. World Scientific Series on Applied Mathematics, 1999.

[18] D. Z. Du and F. Hwang, *Pooling Designs and Nonadaptive Group Testing*. World Scientific Series on Applied Mathematics, 2006.

[19] G. Atia and V. Saligrama, "Boolean compressed sensing and noisy group testing." Manuscript (arXiv: cs.IT/0907.1061)., 2009.

[20] A. Kahng and S. Reda, "Combinatorial group testing methods for the bist diagnosis problem," in *Proceedings of the 2004 Asia and South Pacific Design Automation Conference*, pp. 113–116, 2004.

[21] A. Dyachkov, V. Rykov, and A. Rashad, "Bound of the length of disjunct codes," *Problems of Information Transmission*, vol. 18, pp. 7–13, 1982.

[22] A. Dyachkov and V. Rykov, "A survey of superimposed code theory," *Problem of Control and Information Theory*, vol. 12, pp. 1–13, 1983.

[23] P. Erdos, P. Frankl, and Z. Furedi, "Families of finite sets in which no set is covered by the union of two others," *J. Combin. Theory Ser. A*, vol. 33, pp. 158–166, 1982.

[24] P. Erdos, P. Frankl, and Z. Furedi, "Families of finite sets in which no set is covered by the union of r others," *Israel J. Math*, vol. 51, pp. 79–89, 1985.

[25] F. Hwang and V. Sos, "Non-adaptive hypergeometric group testing," *Studia Sci. Math.*, vol. 22, pp. 257–263, 1985.

[26] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, pp. 489–509, Feb. 2006.

[27] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, Apr. 2006.

[28] A. C. Gilbert and M. J. Strauss, "Group testing in statistical signal recovery," *Technometrics*, vol. 49, no. 3, pp. 346–356, 2007.

[29] A. Gilbert, M. A. Iwen, and M. Strauss, "Group testing and sparse signal recovery," in *Proc. 42nd Asilomar Conference on Signals, Systems and Computers*, 2008.

[30] M. Herman and T. Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *IEEE Journal on Selected Areas in Communications*, vol. 4, no. 2, pp. 342–349, 2010.

[31] M. Akçakaya and V. Tarokh, "Shannon theoretic limits on noisy compressive sampling," *IEEE Transactions on Information Theory*, vol. 56, pp. 492–504, Jan. 2010.

[32] W. Wang, M. J. Wainwright, and K. Ramchandran, "Information-theoretic limits on sparse signal recovery: Dense versus sparse measurement matrices," *IEEE Transactions on Information Theory*, vol. 56, pp. 2946–2979, Jun. 2010.

[33] V. G. A.K. Fletcher, S. Rangan, "Necessary and sufficient conditions on sparsity pattern recovery," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 5758–5772, 2009.

[34] G. Reeves and M. Gastpar, "Sampling bounds for sparse support recovery in the presence of noise," in *IEEE International Symposium on Information Theory*, pp. 2187–2191, July 2008.

[35] F. Hwang, "Group testing with a dilution effect," *Biometrika*, vol. 63, no. 3, pp. 671–673, 1976.

[36] A. Clementi, A. Monti, and R. Silvestri, "Selective families, superimposed codes, and broadcasting on unknown radio networks," in *Proceedings of the 12th Symposium on Discrete Algorithms (SODA)*, 2001.

[37] M. Iwen, "A deterministic sub-linear time sparse fourier algorithm via non-adaptive compressed sensing methods," in *Proceedings of the 19th Symposium on Discrete Algorithms (SODA)*, 2008.

[38] M. Mitzenmacher and E. Upfal, *Probability and Computing*. Cambridge University Press, 2005.

Mahdi Cheraghchi received the B.Sc. degree in computer engineering from Sharif University of Technology, Tehran, Iran, in 2004 and the M.Sc. and Ph.D. degrees in computer science from EPFL, Lausanne, Switzerland, in 2005 and 2010, respectively. Since October 2010, he has been a post-doctoral researcher at the University of Texas at Austin, TX. His research interests include the interconnections between coding theory and theoretical computer science, derandomization theory and explicit constructions.

Ali Hormati received the B.Sc. and M.Sc. degrees in communication systems from Sharif University of Technology, Tehran, Iran, in 2001 and 2003, respectively, and the Ph.D. degree in electrical engineering from EPFL, Lausanne, Switzerland in 2010. From 2003 to 2006, he worked as a communication systems engineer in industry on the communication protocols for software radio systems. His collaboration with Qualcomm company during the Ph.D. studies resulted in patent filings on new techniques for MIMO channel estimation in OFDM and CDMA communication systems. He is now a postdoctoral researcher in the Audiovisual Communications Laboratory at EPFL. His research interests include sampling theory, compressed sensing and inverse problems in acoustic tomography.

Amin Karbasi received the B.Sc. degree in electrical engineering in 2004 and M.Sc. degree in communication systems in 2007 from EPFL , Lausanne, Switzerland. Since March 2008, he has been a Ph.D. student at EPFL. He was the recipient of the ICASSP 2011 best student paper award and ACM/Sigmetrics 2010 best student paper award. His research interests include graphical models, large scale networks, compressed sensing and information theory.

Martin Vetterli received the Dipl. El.-Ing. degree from ETHZ, Switzerland, in 1981, the M.Sc. degree from Stanford University, CA, in 1982, and the Doctorat ès Sciences degree from EPFL, Lausanne, Switzerland, in 1986. He was a research assistant at Stanford and EPFL, and has worked for Siemens and AT&T Bell Laboratories. In 1986 he joined Columbia University in New York, where he was last an Associate Professor of Electrical Engineering and codirector of the Image and Advanced Television Laboratory. In 1993, he joined the University of California at Berkeley, where he was a Professor in the Department of Electrical Engineering and Computer Sciences until 1997, and now holds an Adjunct Professor position. Since 1995 he is a Professor of Communication Systems at EPFL, Switzerland, where he chaired the Communications Systems Division (1996/97), and heads the Audiovisual Communications Laboratory. He has held visiting positions at ETHZ (1990) and Stanford (1998), is a fellow of IEEE, a fellow of ACM, a fellow of EURASIP, and a member of SIAM. He received the Best Paper Award of EURASIP in 1984, the Research Prize of the Brown Bovery Corporation (Switzerland) in 1986, the IEEE Signal Processing Society's Senior Paper Awards in 1991, in 1996 and in 2006 (for papers with D. LeGall, K. Ramchandran, and Marziliano and Blu, respectively). He won the Swiss National Latsis Prize in 1996, the SPIE Presidential Award in 1999, the IEEE Signal Processing Technical Achievement Award in 2001, and the IEEE Signal Processing Society Award in 2010. He has published about 145 journal papers on a variety of topics in signal/image processing and communications, co-authored three books, and holds a dozen patents. His research interests include sampling, wavelets, multirate signal processing, computational complexity, signal processing for communications, digital image/video processing, joint source/channel coding, signal processing for sensor networks and inverse problems like acoustic tomography.