IMPROVEMENT OF ROBUSTNESS TO CHANGE OF POSITIVE ELEMENTS IN BOOLEAN COMPRESSIVE SENSING

Yohei Kawaguchi, Tatsuhiko Osa, Hisashi Nagano, and Masahito Togami

Central Research Laboratory, Hitachi, Ltd.
1-280, Higashi-koigakubo Kokubunji, Tokyo 185-8601, Japan
yohei.kawaguchi.xk@hitachi.com

ABSTRACT

A new boolean compressive sensing method for solving the group-testing problem is proposed. The conventional method has the problem that the estimation performance is degraded in the case that positive elements change in the middle of tests because the results of the tests before a change-point are inconsistent with those of the tests after the change-point. To solve the problem, the proposed method detects the latest change-point of positive elements, and it finds positive elements by using only the results of the tests after the change-point. To detect the change-point, the proposed method makes use of the fact that the distribution of the results depends on the number of positive elements. Experimental simulation indicates that the proposed method outperforms the conventional method on the condition that positive elements change in the middle of tests.

Index Terms— group testing, compressive sensing, change-point detection

1. INTRODUCTION

Group testing is the well-known problem that attempts to discover a sparse subset of positive elements in a large set of elements by using a small number of tests. Each test consists of three processing steps: (1) selecting elements for a pool on the basis of a certain method, (2) mixing the selected elements into the pool, and (3) observing a single Boolean result by testing the pool. When the proportion of positive elements is small, a small number of the tests on the mixed pool are sufficient to detect the positive elements; that is, all the elements need not be tested directly. Group testing as a subject dates back to the work of Dorfman [1] in 1943, during the Second World War. Dorfman developed this approach in order to test soldiers' blood for syphilis. Group testing has applications such as blood screening, deoxyribonucleic acid (DNA) sequencing, and anomaly detection in computer networks [2]. To protect the security in wide areas such as stations, airports, etc., the authors have proposed a system for finding the locations where dangerous substances exist in a short time [3]. The proposed system is based on group testing by using a number of pipes installed around the area and a single mass spectrometer. Air samples are taken from multiple pipes at the same time using a different combination of pipes each time, the mixed samples are analyzed by the mass spectrometer, and the locations of dangerous substances is estimated from the time series of the mass spectrometry signal by group testing. The proposed approach can achieve rapid detection of substances in the area being monitored without requiring a large number of expensive mass spectrometers. A new method for practical application of the system is proposed in this paper.

Traditionally, group testing has been regarded as a combinatorial problem. As for this problem, many researches about the upper and lower bounds on the number of tests required to find all the positive elements have been done. A set of information-theoretic bounds for group testing with random mixing was established by Malyutov [4, 5], Atia and Saligrama [6], Sejdinovic and Johnson [7], and Aldridge *et al.* [8]. In addition, several tractable approximation algorithms, such as one based on belief propagation [7] and one based on matching pursuit [9], have been proposed.

In recent years, group testing has drawn interest from the active research area of compressive sensing. Compressive sensing solves a kind of underdetermined linear equation, namely, $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{x}$ is an unknown high-dimensional vector to be estimated, $\boldsymbol{A}$ is a given mixing matrix, and $\boldsymbol{y}$ is a given low-dimensional observed vector. The problem with compressive sensing is similar to that with group testing from the viewpoint that both of them are underdetermined problems such that an unknown high-dimensional vector is decoded from an observed low-dimensional vector. However, while compressive sensing is defined in a real vector space, group testing is defined in a Boolean vector space. To improve the performance of group testing by using compressive sensing, Malioutov and Malyutov [10] proposed a method for converting group testing into compressive sensing through linear-programming relaxation. As for this conversion method, ℓ_1 minimization imposes the sparsity constraint to the solution and solves the uncertainty of the underdetermined problem. It thus outperforms other conventional methods (i.e., the method based on belief propagation [7], the method based on matching pursuit [9], etc.). In addition, to solve the problem that the number of positive elements is unknown in advance, some approaches changing the mixing matrix adaptively, called adaptive group testing [11], have been proposed. However, the conventional methods assume that positive elements does not change in the middle of tests, and the estimation performance is degraded in the case that positive elements change because the results of the tests before a change-point are inconsistent with those of the tests after the change-point. Also, in the application for location estimation of substances, the locations of substances change in the middle of tests.

To improve the robustness to change of positive elements, a method for group testing is proposed here. The proposed method detects the latest change-point of positive elements, and it finds positive elements by using only the results of the tests after the change-point. To detect the change-point, the proposed method makes use of the fact that the distribution of the results depends on the number of positive elements. Experimental simulation indicates that the proposed method outperforms the conventional method [10] on the condition that positive elements change in the middle of tests.

2. PROBLEM STATEMENT

To state the problem, first, the following notation is fixed. N is the number of elements, of which a subset of size K is positive. Non-positive elements are called negative. $x_n = 1$ indicates that the n-th element is positive, and $x_n = 0$ indicates that the n-th element is negative. For convenience, $\boldsymbol{x} = [x_1, x_2, \cdots, x_N]^T$ is written. T tests, where $T < N$, are then performed. As explained above, in each test, some elements are selected from all the elements, and they are mixed into the same pool. This selection is defined by a mixing matrix, $\boldsymbol{A}$, which is a $T \times N$ binary matrix. The element of the t-th row and the n-th column of $\boldsymbol{A}$ is given as a_{tn}, where $a_{t,n} = 1$ indicates that the n-th element is mixed into the pool of the t-th test, and $a_{t,n} = 0$ indicates that the n-th element is not mixed into the pool of the t-th test. The observed signal of each test, t, is a single Boolean value, $y_t \in \{0, 1\}$. y_t is obtained by taking the Boolean sum of $\{x_n | a_{tn} = 1\}$. For convenience, $\boldsymbol{y} = [y_1, y_2, \cdots, y_T]^T$ is written. The vector notation

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{1}$$

is used in the following.

The problem of group testing is to estimate unknown vector $\boldsymbol{x}$ from given $\boldsymbol{A}$ and $\boldsymbol{y}$. In addition, the noise of the observation is considered. The noise includes both the false positive and the false negative. The former represents the case that $y_t = 1$ even when the Boolean sum of $\{x_n | a_{tn} = 1\}$ is 0. The latter represents the case that $y_t = 0$ even when the Boolean sum of $\{x_n | a_{tn} = 1\}$ is 1. This observation with noise is represented by

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \otimes \boldsymbol{v}, \tag{2}$$

where $\boldsymbol{v}$ is the Boolean vector of errors, and $\otimes$ means the XOR operation.

A number of works have studied the design of $\boldsymbol{A}$ [2]. For example, K-separating and K-disjunct are well-known properties of $\boldsymbol{A}$. When these properties hold, $\boldsymbol{x}$ can be recovered exactly. However, such design is often unsuitable for practical situations because it assumes that the exact number of the positive elements (K) is necessary before group testing. Moreover, if all T tests cannot be carried out, the performance of the method will not be guaranteed [8]. Therefore, in many works, $\boldsymbol{A}$ is simply designed by the Bernoulli random design, where each element of $\boldsymbol{A}$ is generated independently at random with a probability p corresponding with the size of the pool. That is, a_{tn} is 1 with probability p, and a_{tn} is 0 with probability $1-p$. Bernoulli random design is also used in this study.

In a number of applications such as location estimation of substances, the unknown vector $\boldsymbol{x}$ may change in the middle of tests, particularly, $\boldsymbol{x}$ may change at an occasional time point, in other words, a "change-point". One of the problems of group testing is that, on the condition that the unknown vector $\boldsymbol{x}$ changes at the change-point, the elements of $\boldsymbol{y}$ corresponding to the tests before the change-point are inconsistent with those corresponding to the tests after the change-point, and $\boldsymbol{x}$ can not be estimated accurately. The present study thus focuses on an method of the improvement of robustness against change of $\boldsymbol{x}$. Here, we assume that all the tests have a particular order relation of the time t, and $\boldsymbol{x}$ can be rewritten as $\boldsymbol{x}(t)$ considering change of $\boldsymbol{x}(t)$. The task is to estimate the current unknown vector $\boldsymbol{x}(T)$ from $\boldsymbol{A}$ and $\boldsymbol{y}$.

3. BOOLEAN COMPRESSIVE SENSING FOR GROUP TESTING

3.1. Compressive sensing

Malioutov and Malyutov [10] proposed a conversion of group testing into compressive sensing through a linear-programming relaxation. This conventional method is the basis of our method, which is explained in this section.

Many works on compressive sensing have been reported [12]. In this study, a sparse signal, $\boldsymbol{x} \in \mathbb{R}^N$, is assumed, and it is estimated from T measurements $\boldsymbol{y} \in \mathbb{R}^T$ by using a random measurement matrix $\boldsymbol{A}$, where $T < N$. Compressive sensing, namely, decoding $\boldsymbol{x}$, uses the following ℓ_0 minimization:

$$\min_{\boldsymbol{x}} |\boldsymbol{x}|_0 \quad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}. \tag{3}$$

However, Eq. (3) is a NP-hard problem, which cannot be solved practically. Candes et al. [12] proved that if certain conditions hold, $\boldsymbol{x}$ can be decoded exactly by the following ℓ_1 minimization:

$$\min_{\boldsymbol{x}} |\boldsymbol{x}|_1 \quad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}. \tag{4}$$

Since ℓ_1 minimization is a simple linear-programming problem, a number of practicable algorithms can be used to solve it.

3.2. Noise-free case

Equation (1) is similar to constraint equation (4). However, it is not a linear equation in a real vector space but a Boolean equation. It is shown in [10] that (1) can be replaced with a closely related linear formulation: $\mathbf{1} \leq \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}$, and $\mathbf{0} = \boldsymbol{A}_{\mathcal{J}}\boldsymbol{x}$, where $\mathcal{I} = \{t|y_t = 1\}$ is the set of positive test results, and $\mathcal{J} = \{t|y_t = 0\}$ is the set of negative test results. A linear-programming formulation similar to Eq. (4) is therefore given as

$$\begin{aligned} &\min_{\boldsymbol{x}} \left\{ \sum_n x_n \right\} \\ &\text{subject to} \quad \mathbf{0} \leq \boldsymbol{x} \leq \mathbf{1}, \\ &\boldsymbol{A}_{\mathcal{I}}\boldsymbol{x} \geq \mathbf{1}, \quad \boldsymbol{A}_{\mathcal{J}}\boldsymbol{x} = \mathbf{0} \end{aligned} \tag{5}$$

3.3. Noisy case

Because (5) does not model noisy cases, the performance of the method is degraded in noisy cases. One version of [10]'s method thus covers the noisy case by adding slack variables as follows:

$$\begin{aligned} &\min_{\boldsymbol{x},\boldsymbol{\xi}} \left\{ \sum_n x_n + \alpha \sum_t \xi_t \right\} \\ &\text{subject to} \quad \mathbf{0} \leq \boldsymbol{x} \leq \mathbf{1}, \quad \mathbf{0} \leq \boldsymbol{\xi}_{\mathcal{I}} \leq \mathbf{1}, \quad \mathbf{0} \leq \boldsymbol{\xi}_{\mathcal{J}}, \\ &\boldsymbol{A}_{\mathcal{I}}\boldsymbol{x} + \boldsymbol{\xi}_{\mathcal{I}} \geq \mathbf{1}, \quad \boldsymbol{A}_{\mathcal{J}}\boldsymbol{x} = \boldsymbol{\xi}_{\mathcal{J}}, \end{aligned} \tag{6}$$

where $\boldsymbol{\xi} = [\xi_1, \cdots, \xi_T]$ is the vector composed of the slack variables, and α is the regularization parameter that balances the amount of noise and the sparsity of the solution.

4. PROPOSED METHOD

The conventional method assumes that $\boldsymbol{x}$ does not change by time, so the current $\boldsymbol{x}$, i.e., $\boldsymbol{x}(T)$, can not be estimated accurately. To improve the robustness against change of $\boldsymbol{x}(t)$, the proposed method detects the latest change-point $c(T)$, and estimates $\boldsymbol{x}(T)$ using only the results of only the tests after $c(T)$. Here, $c(T)$ is defined as t such that $\boldsymbol{x}(t-1) \neq \boldsymbol{x}(t)$, and $\boldsymbol{x}(t) = \boldsymbol{x}(t+1) = \cdots = \boldsymbol{x}(T)$. To detect the change-point $c(T)$, the proposed method performs a likelihood ratio test as follows:

$$\frac{P(H_1; y_t, \cdots, y_T)}{P(H_0; y_t, \cdots, y_T)} = \frac{1 - P(H_0; y_t, \cdots, y_T)}{P(H_0; y_t, \cdots, y_T)} > \theta \tag{7}$$

where H_0 is the hypothesis that there is no change-point in the tests of $\tau = t \cdots T$, H_1 is the hypothesis that there is a change-point in the tests of $\tau = t \cdots T$, and θ is a threshold parameter. $P(H_0; y_t, \cdots, y_T)$ can be rewritten as:

$$\begin{aligned} & P(H_0; y_t, \cdots, y_T) \\ \propto\ & P(H_0, y_t, \cdots, y_T) \\ =\ & P(\boldsymbol{x}(t) = \cdots = \boldsymbol{x}(T), y_t, ..., y_T) \\ \leq\ & P(|\boldsymbol{x}(t)|_0 = \cdots = |\boldsymbol{x}(T)|_0, y_t, ..., y_T). \end{aligned} \tag{8}$$

In the application of location estimation of substances, not only the combination of positive elements but also the number of positive elements changes at most change-points because the change of $\boldsymbol{x}$ is caused by substance diffusion. Therefore, we approximate (8) as follows:

$$\begin{aligned} & P(\boldsymbol{x}(t) = \cdots = \boldsymbol{x}(T), y_t, ..., y_T) \\ \approx\ & P(|\boldsymbol{x}(t)|_0 = \cdots = |\boldsymbol{x}(T)|_0, y_t, ..., y_T) \end{aligned} \tag{9}$$

Based on the approximation of (9), the proposed method can make use of the fact that the distribution of $\boldsymbol{y}$ depends on $|\boldsymbol{x}(t)|_0$. (9) is converted as follows:

$$\begin{aligned} & P(|\boldsymbol{x}(t)|_0 = \cdots = |\boldsymbol{x}(T)|_0, y_t, ..., y_T) \\ =\ & \sum_K P(|\boldsymbol{x}(t)|_0 = \cdots = |\boldsymbol{x}(T)|_0 = K, y_t, ..., y_T) \\ =\ & \sum_K \prod_{\tau=t}^{T} P(|\boldsymbol{x}(\tau)|_0 = K) \\ & \times \prod_{\tau=t}^{T} P(y_\tau; |\boldsymbol{x}(\tau)|_0 = K). \end{aligned} \tag{10}$$

Here, we can assume that $P(|\boldsymbol{x}(\tau)|_0 = K)$ is a sparse prior distribution, for example, the Poisson distribution $P(|\boldsymbol{x}(\tau)|_0 = K) = \frac{\lambda^K e^{-\lambda}}{K!}$, where λ is the parameter of the distribution. Furthermore, $P(y_\tau; |\boldsymbol{x}(\tau)|_0 = K)$ can be estimated by summation of y_τ, $Y(\tau) = \sum_{f=\tau-F}^{\tau+F} y_f$, which is the sufficient statistics of the Bernoulli distribution, as follows:

$$\begin{aligned} & P(y_\tau; |\boldsymbol{x}(\tau)|_0 = K) \\ =\ & \left\{ 1 - \left(1 - \frac{K}{N} \frac{\sum_{n=1}^{N} a_{\tau n}}{N} \right)^N \right\}^{Y(\tau)} \\ & \times \left\{ \left(1 - \frac{K}{N} \frac{\sum_{n=1}^{N} a_{\tau n}}{N} \right)^N \right\}^{2F+1-Y(\tau)} \end{aligned} \tag{11}$$

where F is the frame size for summation of y_τ. Thus, each time the result of the test is obtained, the proposed method calculates (11), (10), and (8), and it evaluates (7) of each t, and it regards the latest test satisfying (7) as the latest change-point $c(T)$.

Finally, we rewrite (6) as follows:

$$\min_{\boldsymbol{x},\boldsymbol{\xi}} \left\{ \sum_n x_n + \alpha \sum_{t=c(T)}^{T} \xi_t \right\}$$

$$\text{subject to} \quad \mathbf{0} \le \boldsymbol{x} \le \mathbf{1}, \quad \mathbf{0} \le \boldsymbol{\xi}_{\mathcal{I}_c} \le \mathbf{1}, \quad \mathbf{0} \le \boldsymbol{\xi}_{\mathcal{J}_c},$$
$$\boldsymbol{A}_{\mathcal{I}_c}\boldsymbol{x} + \boldsymbol{\xi}_{\mathcal{I}_c} \ge \mathbf{1}, \quad \boldsymbol{A}_{\mathcal{J}_c}\boldsymbol{x} = \boldsymbol{\xi}_{\mathcal{J}_c}, \tag{12}$$

where c is the latest change-point, and $\mathcal{I}_c = \{t|y_t = 1, c(T) < t\}$ is the set of positive test results, and $\mathcal{J}_c = \{t|y_t = 0, c(T) < t\}$ is the set of negative test results. By calculating (12), the proposed method estimate the $\boldsymbol{x}(T)$.

5. EXPERIMENTAL RESULTS

The performance of the proposed method was evaluated by simulation. In particular, the averaged probability of correct estimation was computed over 100 trials as a function of T, for $N = 150$. N elements were generated independently for each trial. In this simulation, the probability p of the Bernoulli random design of $\boldsymbol{A}$ was 0.333, noise with i.i.d 3% probability of flipping each bit of $\boldsymbol{y}$ was added, λ was 1.5, F was 15, and α was 1.0. The proposed method was compared with the conventional method [10]. To evaluate the robustness against change of the unknown vector $\boldsymbol{x}(t)$, the simulation was conducted for three cases:

Case1 $\boldsymbol{x}(t)$ changed from $|\boldsymbol{x}(\tau)|_0 = 0$ to $|\boldsymbol{x}(\tau)|_0 = 2$.

Case2 $\boldsymbol{x}(t)$ changed from $|\boldsymbol{x}(\tau)|_0 = 1$ to $|\boldsymbol{x}(\tau)|_0 = 4$.

Case3 $\boldsymbol{x}(t)$ changed from $|\boldsymbol{x}(\tau)|_0 = 4$ to $|\boldsymbol{x}(\tau)|_0 = 1$.

In all the cases, the change-point, c was 100. As for the conventional method, the latest 20, 50, and 100 tests were used, and all the T tests were used. As for the proposed method, at each test, $c(T)$ was estimated, and the latest $(T - c(T))$ tests were used.

The performance of the proposed method in the case of no noise was computed. Figure 1 shows the probability of exact recovery in **Case1**, Figure 2 shows that in **Case2**, and Figure 3 shows that in **Case3**. "20 TESTS", "50 TESTS", "100 TESTS", and "ALL TESTS" mean respectively the case that the latest 20 tests were used, the case that the latest 50 tests were used, the case that the latest 100 tests were used, and the case that all the tests were used in the conventional method [10]. "PROPOSED" means the proposed method.

First, these results show that the probability of exact recovery of "ALL TESTS" did not increase along with the increase of number of tests after the change-point $c = 100$. This indicates that the case that all the tests were used in the conventional method is not robust against change of positive elements. In all the cases, the probability of exact recovery of the proposed method, i.e., "PROPOSED", converged to 1. In contrast, that of "20 TESTS" in **Case1**, that of "50 TESTS" in

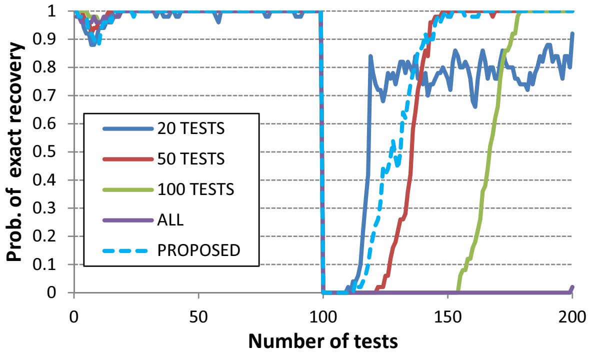

Fig. 1. Probability of exact recovery as a function of number of tests, T, in the case that $\boldsymbol{x}(t)$ changed from $|\boldsymbol{x}(\tau)|_0 = 0$ to $|\boldsymbol{x}(\tau)|_0 = 2$. $N = 150$, the change-point c was 100, and 3% noise was added.

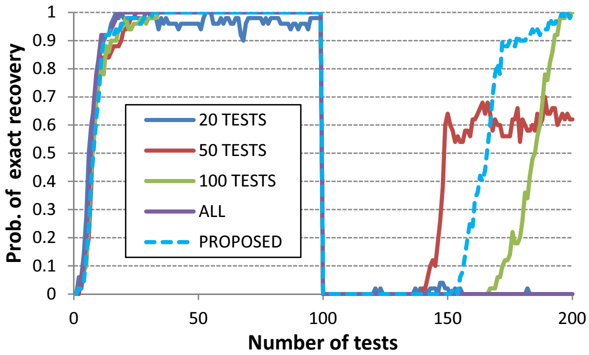

Fig. 2. Probability of exact recovery as a function of number of tests, T, in the case that $\boldsymbol{x}(t)$ changed from $|\boldsymbol{x}(\tau)|_0 = 1$ to $|\boldsymbol{x}(\tau)|_0 = 4$. $N = 150$, the change-point c was 100, and 3% noise was added.

Case2, and that of "50 TESTS" in **Case3** converged to a value lower than 1. These results represent that the fixed numbers of the tests that were used were too small in these cases. In all the cases, the probability of exact recovery of "100 TESTS" converged to 1, however, the speed of the convergence was slower than that of the proposed method. These results represent that the fixed number of the tests that were used was too large in these cases. Thus, it is indicated that the proposed method is robust against change of positive elements.

6. CONCLUSION

A new method for solving the group-testing problem is proposed. To improve the robustness to the condition that positive elements change in the middle tests, the proposed method detects the latest change-point of positive elements, and it finds positive elements by using only the results of the tests after the change-point. To detect the change-point, the proposed

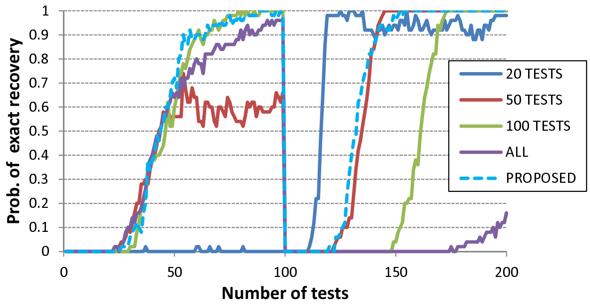

Fig. 3. Probability of exact recovery as a function of number of tests, T, in the case that $\boldsymbol{x}(t)$ changed from $|\boldsymbol{x}(\tau)|_0 = 4$ to $|\boldsymbol{x}(\tau)|_0 = 1$. $N = 150$, the change-point c was 100, and 3% noise was added.

method makes use of the fact that the distribution of the results depends on the number of positive elements. An experimental simulation showed that the proposed method outperforms the conventional method on the condition that positive elements change in the middle tests.

REFERENCES

[1] R. Dorfman, "The detection of defective members of large populations," *Annals of Mathematical Statistics*, vol. 14, no. 6, pp. 436–440, 1943.

[2] D.Z. Du and F.K. Hwang, *Pooling designs and non-adaptive group testing: Important tools for DNA sequencing*, World Scientific, 2006.

[3] T. Kagehiro, K. Yoneji, H. Kiyomizu, Y. Watanabe, Y. Kawaguchi, Z. Li, H. Nagano, and Y. Matsuda, "Traceable physical security systems for a safe and secure society," *Hitachi Review*, vol. 63, no. 5, pp. 254–258, 2014.

[4] M.B. Malyutov, "On planning of screening experiments," in *1975 IEEE-USSR Workshop on Information Theory*, 1976, pp. 144–147.

[5] M.B. Malyutov, "The separating property of random matrices," *Matematicheskie Zametki*, vol. 23, pp. 155–167, 1978.

[6] G. Atia and V. Saligrama, "Boolean compressed sensing and noisy group testing," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1880–1901, 2012.

[7] D. Sejdinovic and O. Johnson, "Note on noisy group testing: Asymptotic bounds and belief propagation reconstruction," in *Allerton conference on communication, control and computing*, 2010.

[8] M. Aldridge and L. Baldassini and O. Johnson, "Group testing algorithms: Bounds and simulations," in *arXiv:1306.6438*, 2013.

[9] C.L. Chan, P.K. Che, S. Jaggi, and V. Saligrama, "Non-adaptive probabilistic group testing with noisy measurements: Near-optimal bounds with efficient algorithms," in *arXiv:1107.4540*, 2011.

[10] D. Malioutov and M. Malyutov, "Boolean compressed sensing: LP relaxation for group testing," in *ICASSP*, 2012.

[11] Y. Kawaguchi, T. Osa, S. Barnwal, H. Nagano, and M. Togami, "Information-based pool size control of boolean compressive sensing for adaptive group testing," in *EUSIPCO*, 2014.

[12] E. Candes and M.B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.