Error Correction Codes for Testing Chemical/Biological Agents: Using Pooled Testing to Increase Test Reliability

Jirong Yi
Dept. of ECE
University of Iowa
Iowa City, IA, 52242, USA
jirong-yi@uiowa.edu

Myung Cho
Dept. of ECE
Penn State Behrend
Erie, PA, 16563, USA
mxc6077@psu.edu

Jingchao Gao
Dept. of Mathematics
University of Iowa
Iowa City, IA, 52242, USA
jingchao-gao@uiowa.edu

Xiaodong Wu
Dept. of ECE
University of Iowa
Iowa City, IA, 52242, USA
xiaodong-wu@uiowa.edu

Weiyu Xu
Dept. of ECE
University of Iowa
Iowa City, IA, 52242, USA
weiyu-xu@uiowa.edu

Raghu Mudumbai
Dept. of ECE
University of Iowa
Iowa City, IA, 52242, USA
rmudumbai@engineering.uiowa.edu

***Abstract*—In this paper, we propose a novel method of error correction code that increases the reliability of testing for chemical or biological agents, as well as general substances, through pooled testing. When gross errors, or outliers, occur in testing results, instead of conducting multiple tests on each individual sample, our method performs testing on well-designed mixtures of samples from multiple subjects. Unlike group testing, which aims to reduce the number of tests, our method aims to increase the reliability and accuracy of testing, even in the presence of gross test errors. Through our theoretical results and extensive simulations, we demonstrate that our method can significantly improve testing accuracy, even under gross testing errors. Furthermore, our method is proven to be more effective than repeated testing of individual samples.**

***Index Terms*—error correction code, non-negative signal, robust testing, pooled testing, test reliability**

I. Introduction

Testing for the presence of chemical or biological agents is crucial in a variety of fields such as science, engineering, and medicine. One common example is the use of Reverse Transcription Polymerase Chain Reaction (RT-PCR) tests to determine whether patient specimens are infected with viruses in the fight against pandemics like COVID-19 [1–4]. However, two major challenges arise when performing such tests: 1) how to test a large number of specimens efficiently and economically, and 2) how to ensure reliable and accurate testing.

This paper focuses on the challenge of performing reliable tests for specimens, even when gross errors occur in some testing results. Gross errors or outliers can occur due to deficiencies in testing technology, sample contamination, or sample dilution. For instance, if a sample is contaminated, the testing result for that sample may change from negative to positive, falsely indicating the presence of non-existent chemical or biological agents.

One common method for dealing with gross errors in testing is to repeat the testing on a single sample several times in the hopes of obtaining accurate results. However, this approach is highly inefficient and expensive, particularly when testing a large number of specimens. In this paper, we propose a novel method of error correction code to increase the reliability of testing for chemical or biological agents, as well as general substances, via pooled testing. Our approach involves pooling specimens into well-designed groups and testing for the targeted chemical or biological agents in these pooled samples. While outliers or gross errors can still occur in individual pools, the pooling strategy serves as an error correction code that can automatically correct for such errors. Moreover, our proposed method is more powerful and efficient than repeated testing of individual specimens, allowing for accurate testing results with fewer tests.

Traditional group testing is a well-known technique that uses pooled testing to reduce the number of tests required. Group testing has received a lot of attention during the COVID-19 pandemic due to its high throughput, with pooled samples from multiple test subjects tested collectively rather than individually [5–11]. The principle behind group testing is that a single negative test result on a pooled sample indicates that all subjects in that pool are infection-free. This allows for a reduction in the total number of tests required per subject, thereby increasing testing throughput, particularly in populations where the infection rate (prevalence) is low [12]. However, highly accurate tests are required for group testing to be effective, as a single false negative result could potentially result in incorrect diagnosis. While improved test accuracy can be achieved by testing each sample multiple times, this

weakens the primary purpose of using group testing, which is to reduce the number of tests required.

In contrast to traditional group testing, the pooled sample testing method proposed in this paper serves a different purpose. Rather than reducing the number of tests, it aims to increase the reliability and accuracy of testing, even in the presence of gross errors in the testing results. The proposed method uses ideas from compressed sensing and error correction coding to correct errors in test results. By pooling samples in a well-designed manner, the redundancy of information among pooled sample mixtures can be exploited to correct wrong test results, similar to how error correction codes correct errors in communication channels [13]. This approach is more efficient than simply testing each individual sample multiple times, as demonstrated by simulations and theoretical arguments showing its effectiveness in reducing false positives and false negatives, even with highly error-prone tests.

Our work is related to one of the seminal papers in compressed sensing [14], which uses linear programming to perform decoding under channel errors. However, our approach differs in that we purposefully design binary matrices for pooled testing, which is related to building codebooks in error correction codes, instead of being given a particular channel as in [14]. In addition, we deal with non-negative signal models, which brings additional structure for sensing and inference [15, 16]. Compared to recent works [17–25] that aim to boost test robustness against noisy measurements in group testing and compressed sensing, our work has the following distinguishable contributions in testing chemical or biological agents:

- The main purpose of our proposed pooled testing is to increase test reliability, rather than reducing the required number of tests as in group testing/compressed sensing, although our approach can provide error correction capability even with a reduced number of tests.
- Our pooled testing using error-correcting codes can be conducted not only in the "undersampling" regime but also in the "oversampling" regime, where the number of performed tests is larger than the number of subjects unlike compressed sensing in the "undersampling" regime.
- We do not necessarily require low prevalence or a sparse signal. The signal considered can be fully dense.
- The proposed error correction pooled testing technique is not restricted to virus testing, but it is also applicable to many other areas of measurements for the detection of chemical or biological agents, where there may be outlier errors in measurements.

Our findings challenge the conventional wisdom that pooling samples together would lead to lower test accuracy or reliability compared to individual separate testing, due to factors such as sample dilutions and signal mixing. Instead, our results demonstrate that, in some cases, purposeful pooled testing can significantly increase, rather than decrease, test accuracy or reliability.

Notations: We reserve capital letters, e.g., $\boldsymbol{X}$, small bold letters, e.g., $\boldsymbol{x}$, and non-bold letters, e.g., x or X_{ij}, for matrices, vectors, and scalars respectively. To represent an element in a vector or a matrix, we use a sub-script, e.g., x_i for the i-th element of $\boldsymbol{x}$, and X_{ij} for an element in the i-th row and the j-th column element of $\boldsymbol{X}$. We denote a set of $m \times n$ binary matrices, and a set of $m \times n$ matrices whose elements are between a and b as $\{0,1\}^{m\times n}$ and $[a,b]^{m\times n}$ respectively. A set of real numbers is denoted by $\mathbb{R}$. $\|\boldsymbol{x}\|_0$, $\|\boldsymbol{x}\|_1$, and $\|\boldsymbol{x}\|_2$ represent ℓ_0 norm of $\boldsymbol{x}$, i.e., the number of non-zero elements of $\boldsymbol{x}$, ℓ_1 norm of $\boldsymbol{x}$, i.e., $\|\boldsymbol{x}\|_1 = \sum_{j=1}^{n} |x_j|$ for $\boldsymbol{x} \in \mathbb{R}^n$, and ℓ_2 norm of $\boldsymbol{x}$ respectively. $\boldsymbol{x} \geq 0$ represents element-wise non-negativeness.

II. Problem Formulation

Let us assume that we have n subjects (sources of specimens), and we have a budget of performing m tests to determine the quantities of the target chemical/biological agent in the specimens taken from these subjects. We denote the quantity (density) of the target chemical/biological agent from the n subjects by $\boldsymbol{x} \in [0,\infty)^n$. For each of the m tests, we create a pooled sample by mixing the specimens from multiple subjects. We use a matrix $\boldsymbol{P} \in \{0,1\}^{m\times n}$ to denote the participation of n subjects in m tests, i.e. the sample of the j-th ($1 \leq j \leq n$) subject is involved in the i-th ($1 \leq i \leq m$) test if $P_{ij} = 1$; and the sample of the j-th subject will not be involved in the i-th test if $P_{ij} = 0$. This means that the number of 1's in the j-th column of $\boldsymbol{P}$ is the number of tests that the specimens of j-th subject is involved in. We model the amount of the subjects' specimens by a matrix $\boldsymbol{W} \in \mathbb{R}^{m\times n}$, and each W_{ij} represents how much of the j-th subject's specimen is used in the i-th test. With those setups, we have a measurement matrix $\boldsymbol{A} := \boldsymbol{P} \odot \boldsymbol{W}$, where $\odot$ represents the Hadamard multiplication. For simplicity of presentation, we assume that $\boldsymbol{W}$ is an all-1 matrix.

The corresponding m mixed samples will go through m quantitative (or sometimes qualitative) tests to test for the target agent. Due to potential background noises and gross errors caused by factors such as dilutions, sample and reagent contamination, and operational mistakes, the final quantitative measurements $\boldsymbol{y} \in \mathbb{R}^m$ from the m tests can be modeled as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}^\star + \boldsymbol{v} + \boldsymbol{e}, \tag{1}$$

where $\boldsymbol{x}^\star \in [0,\infty)^n$, $\boldsymbol{v} \in \mathbb{R}^m$, and $\boldsymbol{e} \in \mathbb{R}^m$ are the ground-truth signal, noises in observation, and possible gross (outlier) errors respectively. Since the elements of the vector $\boldsymbol{x}^\star$ represent the quantities of target agents, the elements of $\boldsymbol{x}^\star$ are non-negative. Then, our goal here is to recover the ground truth signal $\boldsymbol{x}^\star \in [0,\infty)^n$ for n subjects from m test measurements $\boldsymbol{y} \in \mathbb{R}^m$ with possible outliers and noises.

III. Matrix Design and Pooled Testing as Error Correction Codes

In our design for measurement matrix, we pick $\boldsymbol{A}$ as matrix with 0 or 1 as its elements. The matrix $\boldsymbol{A}$ should be designed such that it can correct gross errors in testing results.

Intuitively, we can think of matrix $\boldsymbol{A}$ as corresponding to the generator matrix of an error correction code. The matrix $\boldsymbol{A}$ should be such that there are enough separations in the m testing results for different input vector $\boldsymbol{x}$, leading to detection of gross errors in testing results. With the designed measurement matrix $\boldsymbol{A}$, we can check the recovery performance of the measurement matrix through the methods proposed in [26].

We formulate the problem of recovering $\boldsymbol{x} \in [0,\infty)^n$ from $\boldsymbol{y} \in \mathbb{R}^m$, where m can be bigger than n, as

$$\begin{aligned} &\underset{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{v}\in\mathbb{R}^m}{\text{minimize}} \ \|\boldsymbol{x}\|_0 + \lambda\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}-\boldsymbol{v}\|_0, \\ &\text{subject to } \ \|\boldsymbol{v}\|_2 \le \epsilon, \ \ \boldsymbol{x} \ge 0. \end{aligned} \tag{2}$$

Here $\lambda \ge 0$ is a tuning parameter for controlling the trade-off between $\|\boldsymbol{x}\|_0$ and $\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}-\boldsymbol{v}\|_0$, and $\epsilon \ge 0$ is a parameter controlling the tolerance for the noise. When the signal $\boldsymbol{x}$ is not sparse, one can drop $\|\boldsymbol{x}\|_0$ in the objective function.

Due to the combinatorial complexity of (2), by relaxing $\|\cdot\|_0$ to $\|\cdot\|_1$, we reformulate (2) as

$$\begin{aligned} &\underset{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{v}\in\mathbb{R}^m}{\text{minimize}} \ \|\boldsymbol{x}\|_1 + \lambda\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}-\boldsymbol{v}\|_1, \\ &\text{subject to } \ \|\boldsymbol{v}\|_2 \le \epsilon, \ \ \boldsymbol{x} \ge 0. \end{aligned} \tag{3}$$

We refer to (3) as ℓ_1–ℓ_1 minimization for the error-correction-code-based pooled testing. If $x_j \ge \tau$, where τ is the threshold value, then we claim the j-th subject is infected and tests "positive"; otherwise, we declare "negative" for the subject. The threshold is introduced to reduce the impact of noise may exists after recovery. More specifically, If τ is too high, all results from both individual and pooled testing will be negative. Conversely, if τ is too low, there will be an increase in testing error in both individual and pooled testing. However, the error-correction capability of pooled testing allows it to suppress testing errors as the number of measurements increases, while individual testing is significantly affected by testing errors.

For theoretical analysis, we consider the simplified but essential case of only having gross errors $\boldsymbol{e}$ but not observation noise $\boldsymbol{v}$, namely $\epsilon = 0$. We also consider the case of dense $\boldsymbol{x}$ (namely λ is large, equivalently dropping $\|\boldsymbol{x}\|_1$).

Theorem III.1. *The optimization program* (3) *uniquely recovers the ground-truth signal $\boldsymbol{x}^\star$ under every $\boldsymbol{e}$ with no more than l gross errors, if for every subset $K \subseteq \{1,2,...,m\}$ with cardinality no more than l and every $\boldsymbol{u} \neq 0$, we have $2\|(\boldsymbol{A}\boldsymbol{u})_K\|_1 < \|(\boldsymbol{A}\boldsymbol{u})\|_1$. For a given 0-1 matrix $\boldsymbol{A}$, the number of gross errors, l, that can always be corrected via solving* (3) *is lower bounded by the floor of*

$$(\min_{\|\boldsymbol{u}\|_2=1} \|\boldsymbol{A}\boldsymbol{u}\|_1)/(2\sqrt{n}) - 1. \tag{4}$$

When the elements of $\boldsymbol{A}$ are chosen i.i.d. following 0-1 Bernoulli distribution with parameter 0.5, and $m = O(n\log(n))$, then with high probability, the correctable number l of gross errors scales at least in the order of $\sqrt{n}$.

We give an outline of the proof of Theorem III.1 due to space limitations. The condition $2\|(\boldsymbol{A}\boldsymbol{u})_K\|_1 < \|(\boldsymbol{A}\boldsymbol{u})\|_1$ in Theorem 1 can be proved in the same way as proving the null space condition for compressed sensing, for example, in [27]. Then the lower-bound on the number of correctable errors (l) as shown in (4) is obtained by noticing that each element of $\boldsymbol{A}\boldsymbol{u}$ is no more than $\|\boldsymbol{u}\|_2\sqrt{n}$ in magnitude according to the Cauchy-Schwarz inequality. Under such a number l, $2\|(\boldsymbol{A}\boldsymbol{u})_K\|_1 < \|(\boldsymbol{A}\boldsymbol{u})\|_1$ holds. When $\boldsymbol{A}$ has i.i.d. Bernoulli elements, we can prove $\sqrt{n}$ errors are correctable through an ϵ-net argument. The idea is to bound $(\min_{\|\boldsymbol{u}\|_2=1} \|\boldsymbol{A}\boldsymbol{u}\|_1)$ by first considering a finite number of points $\boldsymbol{u}$ on the unit sphere, called the ϵ-net, and developing concentration of measure inequalities for the ϵ-net before extending the bounding for every $\boldsymbol{u}$ on the unit sphere. This approach was also used in our earlier results for Gaussian matrices $\boldsymbol{A}$ [27].

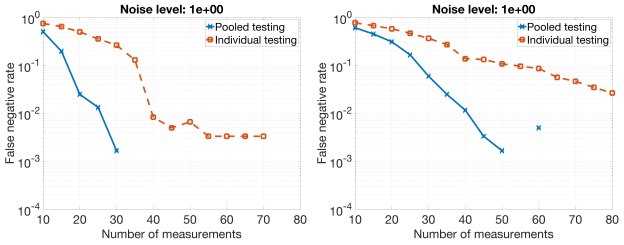
Our analysis shows the pooled method can correct more errors than repeated testings of individual subjects. Consider $\log(n)$ tests are allocated for repeatedly testing each individual subject. Then in individual testing $\sqrt{n}$ testing errors may not be correctable. However, pooled error correction code approach can correct $\sqrt{n}$ (in order) errors.

IV. Numerical Experiments

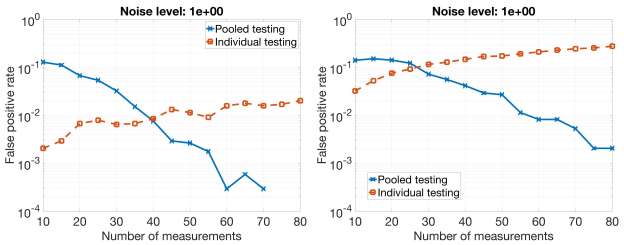
We conduct numerical experiments to evaluate the performance of the proposed pooled testing in Section III. Our method can be applied to general chemical/biological agent testing. We consider that k out of n subjects are "infected" with the target chemical/biological agent. We randomly chose k elements in $\boldsymbol{x}$ to be positive and other $n-k$ elements to zero, where $n = 40$. For the values of non-zero elements in $\boldsymbol{x}$, we choose real numbers between 5 and 10 uniformly at random. The sparsity level k is varied from 3 to 6.

In the pooled testing, the Gaussian noise vector $\boldsymbol{v}$ in (1) is generated i.i.d. following the Gaussian distribution $\mathcal{N}(0,\sigma^2)$, where the noise level σ^2 is varied from $5e-1$ to $2e0$. In order to reflect false positive or false negative error scenarios, we generate our signals as follows. We firstly choose the locations of non-zero elements of the sparse outlier vector $\boldsymbol{e}$ with probability $\mathbb{P}_{out}$. If we have an outlier in the i-th measurement, i.e., y_i, we consider whether $(\boldsymbol{A}\boldsymbol{x})_i$ is positive (infected case) or zero (non-infected case). If $(\boldsymbol{A}\boldsymbol{x})_i$ is positive, with 95% probability, we set the outlier e_i to $-(\boldsymbol{A}\boldsymbol{x})_i$ and reset $v_i = 0$ such that $y_i = 0$, i.e., false negative measurement. With the other 5% probability, we set the outlier e_i to $5\times q+2$, where q follows the standard Gaussian distribution $\mathcal{N}(0,1)$, and keep the originally generated v_i, which is a noise measurement case. If $(\boldsymbol{A}\boldsymbol{x})_i$ is zero, then, we set e_i to $5\times|q|+2$, where q follows distribution $\mathcal{N}(0,1)$, and keep the originally generated v_i, which leads to non-zero measurement y_i, i.e., possibly infected measurement result. Since the measurement $\boldsymbol{y}$ must be a non-negative vector, we have $y_i = \max\{(\boldsymbol{A}\boldsymbol{x})_i+e_i+v_i, 0\}$, $i = 1,...,m$, to prevent y_i being negative. The probability of the outlier error, $\mathbb{P}_{out}$, is varied from 1% to 15%. For the pooled testing, we use (3) to recover $\boldsymbol{x}$, and use threshold $\tau = 1$ to decide whether a subject tests positive or negative.

We compare the pooled testing against the individual testing, where the individual samples of subjects are tested separately

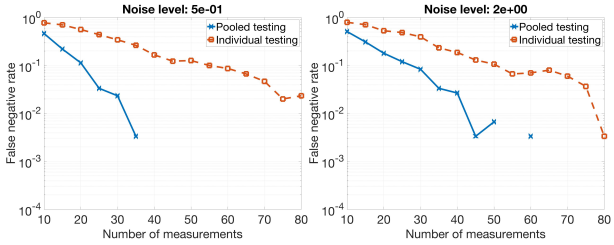


(a) FNR ($\mathbb{P}_{out}$ = 0.01) (b) FNR ($\mathbb{P}_{out}$ = 0.15)

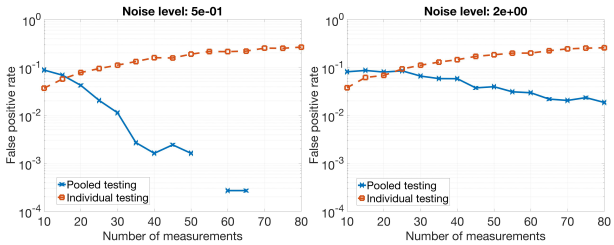


(c) FPR ($\mathbb{P}_{out}$ = 0.01) (d) FPR ($\mathbb{P}_{out}$ = 0.15)

Fig. 1. Simulations for different probabilities of outlier errors. False Negative Rate (FNR) and the corresponding False Positive Rate (FPR) with $n = 40$, $k = 6$, and Gaussian noise level 1e0.



(a) FNR (Noise level: 5e-1) (b) FNR (Noise level: 2e0)



(c) FPR (Noise level: 5e-1) (d) FPR (Noise level: 2e0)

Fig. 2. Simulations for different Gaussian noise levels. False Negative Rate (FNR) and the corresponding False Positive Rate (FPR) with $n = 40$, $k = 3$, $\mathbb{P}_{out} = 0.15$.

(possibly multiple times). In the individual testing, the i-th test is modeled as $y_i = x_{\mod(i-1,n)+1} + e_i + v_i$, $i = 1, 2, ..., m$. We generate the Gaussian noises and outliers in the same way as described for the pooled testing, based on $x_{\mod(i-1,n)+1}$ instead of $(\boldsymbol{Ax})_i$. For example, for $n = 40$, the 42-th measurement y_{42} is the result for the 2nd subject (this subject has been tested once already in the 2nd test), and the outlier and the Gaussian noise simulated for the 42-th test is randomly generated based on x_2. In the individual testing, if $m < n$, there must be some subjects not getting tested at all; we consider the untested subjects as negative in our simulations. Additionally, in the individual testing, if a subject is tested multiple times, and as long as one of the results is identified as being positive, we consider the subject as positive. This comes from the motivation of not missing any positive case (e.g., in COVID-19 virus testing). The number of measurements, m, is varied from 10 to 80. Thus, for $n = 40$ and individual testing scenario, the maximum number of tests for a subject is two.

For both the pooled testing and the individual testing, we run 100 random trials for each parameter setup, and record the False Negative Rate (FNR) and the False Positive Rate (FPR), which are computed on average out of 100 trials as follows:

$$\text{FNR} = \frac{\text{\# of decoded-to-be negative cases in infected subjects}}{\text{\# of subjects truly infected}},$$

$$\text{FPR} = \frac{\text{\# of decoded-to-be positive cases in subjects not infected}}{\text{\# of people not infected}}.$$

Hence, the FNR represents the percentage of subjects decoded as negative among subjects infected with target chemical/biological agent, which can be a critical error in the testing. The FPR is interpreted as the percentage of subjects who are diagnosed as infected among subjects who are in fact not infected. The lower FNR and FPR are, the better testing performance we have.

Figs.1 and 2 show the comparison results between the proposed pooled testing and the individual testing, where the blue solid line and the red dotted line represent the pooled testing and the individual testing respectively. As shown in Figs.1 and 2, the pooled testing lowers both the FNR and the FPR as the number of measurements increases. This is because as the number of measurements increases, we can recover more accurate result $\boldsymbol{x}$ via $\ell_1 - \ell_1$ minimization in (3). Unlike the pooled testing, the individual testing can reduce the FNR as the number of measurements increases at the cost of increasing the FPR. This is because the individual testing diagnoses the subject "positive" once we have one positive test result among multiple tests. Additionally, even without this conservative strategy in the individual testing, when the number of measurements is 40, as in the individual testing method, where each subject is tested once, our pooled testing method has a smaller FNR and FPR as shown in the figures. This suggests that our method has improved reliability compared to the individual testing method. These various simulation results clearly demonstrate that the pooled testing can provide lower FNR and FPR than those of the individual testing. In a limited number of cases, the individual testing provides lower (although not significantly lower) FPR than that of the pooled testing, because only a few subjects are tested under individual testing, which leads to fewer false positive errors. Recall that the untested subjects are assumed to be diagnosed as "negative". Additionally, for $m < n$, since there are simply untested subjects in the individual testing, the individual testing has relatively higher FNR.

REFERENCES

[1] U. E. Gibson, C. A. Heid, and P. M. Williams, "A novel method for real time quantitative RT-PCR," *Genome Research*, vol. 6, no. 10, pp. 995–1001, 1996.

[2] T. Nolan, R. Hands, and S. Bustin, "Quantification of mRNA using real-time RT-PCR," *Nature Protocols*, vol. 1, no. 3, pp. 1559–1582, 2006.

[3] B. D. Kevadiya, J. Machhi, J. Herskovitz, M. D. Oleynikov, W. R. Blomberg, N. Bajwa, D. Soni, S. Das, M. Hasan, M. Patel, et al., "Diagnostics for SARS-CoV-2 infections," *Nature materials*, vol. 20, no. 5, pp. 593–605, 2021.

[4] A. F. Sunjaya and A. P. Sunjaya, "Pooled testing for expanding COVID-19 mass surveillance," *Disaster Medicine and Public Health Preparedness*, vol. 14, no. 3, pp. e42–e43, 2020.

[5] N. Sinnott-Armstrong, D. Klein, and B. Hickey, "Evaluation of group testing for SARS-CoV-2 RNA," *medRxiv 2020.03.27.20043968*, 2020.

[6] H. Shani-Narkiss, O. Gilday, N. Yayon, and I. Landau, "Efficient and practical sample pooling for high-throughput PCR diagnosis of COVID-19," *medRxiv 2020.04.06.20052159*, 2020.

[7] C. Hogan, M. Sahoo, and B. Pinsky, "Sample pooling as a strategy to detect community transmission of SARS-CoV-2," *Journal of the American Medical Association (JAMA)*, vol. 323, no. 19, pp. 1967–1969, 2020.

[8] K. A. Waldstein, J. Yi, M. Cho, R. Mudumbai, X. Wu, S. M. Varga, and W. Xu, "Use of compressed sensing to expedite high-throughput diagnostic testing for COVID-19 and beyond," *PLOS Computational Biology*, vol. 18, no. 10, pp. 1–20, 2022.

[9] I. Yelin, N. Aharony, E. Shaer-Tamar, A. Argoetti, E. Messer, D. Berenbaum, E. Shafran, A. Kuzli, N. Gandali, T. Hashimshony, Y. Mandel-Gutfreund, M. Halberthal, Y. Geffen, M. Szwarcwort-Cohen, and R. Kishony, "Evaluation of COVID-19 RT-qPCR test in multi-sample pools," *medRxiv 2020.03.26.20039438*, 2020.

[10] J. Žilinskas, A. Lančinskas, and M. R. Guarracino, "Pooled testing with replication as a mass testing strategy for the COVID-19 pandemics," *Scientific Reports*, vol. 11, no. 1, pp. 1–7, 2021.

[11] Y.-J. Lin, C.-H. Yu, T.-H. Liu, C.-S. Chang, and W.-T. Chen, "Positively correlated samples save pooled testing costs," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 3, pp. 2170–2182, 2021.

[12] R. Hanel and S. Thurner, "Boosting test-efficiency by pooled testing for SARS-CoV-2formula for optimal pool size," *PLOS ONE*, vol. 15, no. 11, pp. 1–10, 2020.

[13] S. Lin and D. J. Costello, *Error control coding*, vol. 2, Prentice hall New York, 2001.

[14] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[15] M. Khajehnejad, A. Dimakis, W. Xu, and B. Hassibi, "Sparse recovery of nonnegative signals with minimal expansion," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, 2011.

[16] M. Wang, W. Xu, and A. Tang, "A unique nonnegative solution to an underdetermined system: From vectors to matrices," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1007–1016, 2011.

[17] J. Zhu, K. Rivera, and D. Baron, "Noisy pooled PCR for virus testing," *arXiv:2004.02689*, 2020.

[18] H. Petersen, B. Bah, and P. Jung, "Practical high-throughput, non-adaptive and noise-robust SARS-CoV-2 testing," *arXiv:2007.09171*, 2020.

[19] J. Yi, R. Mudumbai, and W. Xu, "Low-cost and high-throughput testing of COVID-19 viruses and antibodies via compressed sensing: system concepts and computational experiments," *arXiv:2004.05759*, 2020.

[20] S. Ghosh, R. Agarwal, M. Rehan, S. Pathak, P. Agrawal, Y. Gupta, S. Consul, N. Gupta, Ritika, R. Goenka, A. Rajwade, and M. Gopalkrishnan, "A compressed sensing approach to pooled RT-PCR testing for COVID-19 detection," *IEEE Open Journal of Signal Processing*, vol. 2, pp. 248–264, 2021.

[21] L. Mutesa, P. Ndishimye, Y. Butera, J. Souopgui, A. Uwineza, R. Rutayisire, E. L. Ndoricimpaye, E. Musoni, N. Rujeni, T. Nyatanyi, et al., "A pooled testing strategy for identifying SARS-CoV-2 at low prevalence," *Nature*, vol. 589, no. 7841, pp. 276–280, 2021.

[22] S. Ghosh, R. Agarwal, M. A. Rehan, S. Pathak, P. Agarwal, Y. Gupta, S. Consul, N. Gupta, R. Goenka, A. Rajwade, et al., "A compressed sensing approach to pooled RT-PCR testing for COVID-19 detection," *IEEE Open Journal of Signal Processing*, vol. 2, pp. 248–264, 2021.

[23] J. Yang and A. Prater-Bennette, "Pooled testing with compressive sensing," in *Proceedings of Big Data III: Learning, Analytics, and Applications*. SPIE, 2021, vol. 11730, pp. 96–103.

[24] A. Cohen, N. Shlezinger, A. Solomon, Y. C. Eldar, and M. Médard, "Multi-level group testing with application to one-shot pooled COVID-19 tests," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 1030–1034.

[25] E. F. Ben-Knaan, Y. C. Eldar, and N. Shlezinger, "Recovery of noisy pooled tests via learned factor graphs with application to COVID-19 testing," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 4518–4522.

[26] M. Cho, K. Vijay Mishra, and W. Xu, "Computable performance guarantees for compressed sensing matrices," *EURASIP journal on advances in signal processing*, vol. 2018, pp. 1–18, 2018.

[27] M. Wang, W. Xu, and A. Tang, "On the performance of sparse recovery via ℓ_p-minimization $(0 \leq p \leq 1)$," *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7255–7278, 2011.