

# Group Testing with Blocks of Positives and Inhibitors

Thach V. Bui <sup>1,\*</sup>, Isao Echizen <sup>2,3</sup>, Minoru Kuribayashi <sup>4</sup>, Tetsuya Kojima <sup>5</sup> and Thuc D. Nguyen <sup>6,7</sup>

<sup>1</sup> Department of Computer Science, National University of Singapore, Singapore 117417, Singapore

<sup>2</sup> National Institute of Informatics, Tokyo 101-8430, Japan

<sup>3</sup> Department of Information and Communication Engineering, University of Tokyo, Tokyo 113-8654, Japan

<sup>4</sup> Graduate School of Natural Science and Technology, Okayama University, Okayama 700-8530, Japan

<sup>5</sup> National Institute of Technology, Tokyo College, Hachioji, Tokyo 193-0997, Japan

<sup>6</sup> Faculty of Information Technology, University of Science, VNU-HCMC, Ho Chi Minh City 72711, Vietnam

<sup>7</sup> Faculty of Information Technology, Vietnam National University, Ho Chi Minh City 720300, Vietnam

\* Correspondence: bvthach@comp.nus.edu.sg

**Abstract:** The main goal of group testing is to identify a small number of specific items among a large population of items. In this paper, we consider specific items as positives and inhibitors and non-specific items as negatives. In particular, we consider a novel model called *group testing with blocks of positives and inhibitors*. A test on a subset of items is positive if the subset contains at least one positive and does not contain any inhibitors, and it is negative otherwise. In this model, the input items are linearly ordered, and the positives and inhibitors are subsets of small blocks (at unknown locations) of consecutive items over that order. We also consider two specific instantiations of this model. The first instantiation is that model that contains a single block of consecutive items consisting of exactly known numbers of positives and inhibitors. The second instantiation is the model that contains a single block of consecutive items containing known numbers of positives and inhibitors. Our contribution is to propose efficient encoding and decoding schemes such that the numbers of tests used to identify only positives or both positives and inhibitors are less than the ones in the state-of-the-art schemes. Moreover, the decoding times mostly scale to the numbers of tests that are significantly smaller than the state-of-the-art ones, which scale to both the number of tests and the number of items.

**Keywords:** non-adaptive group testing; inhibitors; combinatorics; sub-linear algorithms; sparse recovery



**Citation:** Bui, T.V.; Echizen, I.; Kuribayashi, M.; Kojima, T.; Nguyen, T.D. Group Testing with Blocks of Positives and Inhibitors. *Entropy* **2022**, *24*, 1562. <https://doi.org/10.3390/e24111562>

Academic Editors: Sergio Cruces, Iván Durán-Díaz, Rubén Martín-Clemente, Andrzej Cichocki and Alessandro Giuliani

Received: 8 September 2022

Accepted: 27 October 2022

Published: 30 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Group testing [1] was first introduced to reduce time and cost of testing draftees who were possibly positive for syphilis. In this problem, the number of syphilitic draftees is outnumbered by the number of non-syphilitic draftees. The main idea of group testing is instead of testing draftees individually, sets of draftees are pooled and tested. If the test outcome of a pool is positive, then there exists at least one draftee in that pool that is syphilitic and none of the draftees in the pool are syphilitic otherwise. Since this seminal work, group testing has been usually treated as a problem of identifying a small number of specific items in a large population of items. The specific items depend on context and affect how a test on a subset of items is positive or negative.

There are two general strategies for designing tests [2]. The first is adaptive group testing in which the design of a test depends on the designs of the previous tests. This approach usually attains an information-theoretic bound for the number of tests but consumes a substantial amount of time for implementation because of several design stages. To remedy its time-consuming nature while achieving a relatively low number of tests, non-adaptive group testing (NAGT) is used. In this strategy, all tests are designed independently and can be performed in parallel. Because of its advantage, NAGT has been used in a wide range of applications, such as computational and molecular biology [2,3], networking [4],

COVID-19 [5,6], and neuroscience [7]. In this work, our focus is on the non-adaptive testing strategy.

NAGT can be represented by a  $t \times n$  binary matrix  $T = (t_{ij})$ , where  $n$  is the number of items and  $t$  is the number of tests. An entry  $t_{ij} = 1$  means that item (column)  $j$  belongs to test (row)  $i$ , and  $t_{ij} = 0$  means otherwise. The  $j$ th item is represented by the  $j$ th column of the matrix. The procedure to produce the measurement matrix is called *construction*, the procedure to obtain the outcomes of tests using the measurement matrix is called *encoding*, and the procedure to recover specific items from the outcomes is called *decoding*. A measurement matrix is *random* if some tests are generated by a probabilistic scheme, whereas it is *deterministic* if every test is deterministic. A measurement matrix is *strongly explicit* (explicit) if it takes the time and space polynomial of the number rows (respectively, the number rows and the number of columns) to generate a column in it.

Some distribution settings may apply on specific items. There are two common settings: (i) the probabilistic setting, in which there is some probability distribution used on specific items, and the identification error probability is allowed; and (ii) the combinatorial setting, which is our focus here, and no probability distribution is used on specific items.

Consider standard group testing in which specific items are only positives. Suppose a test on a subset of items is positive if the subset contains at least one positive and is negative otherwise. Throughout the paper, log refers to base 2 logarithms. If we give a population of  $n$  items up to  $d$  positives, then there are a number of works for attaining a low number of tests, say  $t = O(d^2 \log^{1+o(1)} n)$ , and/or a fast decoding time, say  $\text{poly}(d, \ln n)$  [8–14] in the combinatorial setting. In probabilistic settings, Bondorf [15] et al. show that the number of tests can be reduced to  $O(d \log n)$  with a decoding time of  $O(d^2 \log d \cdot \log n)$ . Price and Scarlett [16] later improved the decoding time to  $O(d \log n)$ .

### 1.1. New Model and Problem Definition

Because of the natural phenomenon in biology, a new type of item called *inhibitor* was introduced in group testing [3] and studied [17–20]. An inhibitor item causes a negative outcome for any test it is involved in. On the other hand, a test on a subset of items is positive if the subset does not contain any inhibitor and contains at least one positive.

Group testing with blocks of positives has been recently presented by Bui et al. [21], which is a generalization of group testing with consecutive positives [22–27]. In this model, input  $n$  items are linearly ordered, and all positives belong to at most  $k$  blocks of consecutive items and each block has up to  $d$  consecutive items.

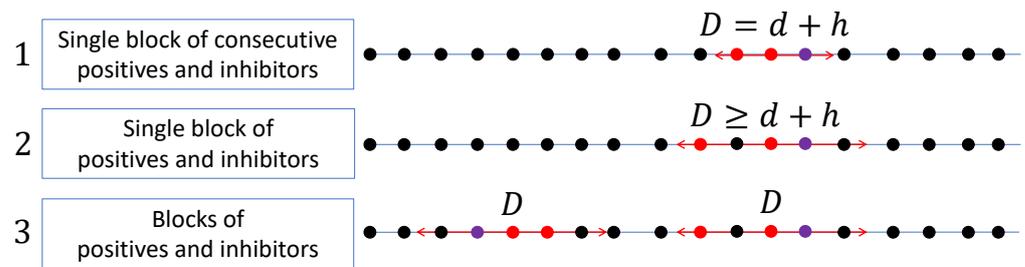
Combining the two models above, we consider a novel model called *group testing with blocks of positives and inhibitors*. The input  $n$  items are linearly ordered. We sub-categorize the model into three models and illustrate them in Figure 1. The first model contains one block of  $d + h$  consecutive items and that block contains exactly  $d$  positives and  $h$  inhibitors. The second model, which is a general model of the first one, contains one block of  $D \geq d + h$  consecutive items and that block contains up to  $d$  positives and  $h$  inhibitors. The third model, which is the most general one, contains multiple blocks, says  $k$ , of consecutive items in which each block of size up to  $D \geq d + h$  contains up to  $d$  positives and  $h$  inhibitors. Note that the assumption on the known upper bounds for  $k$ ,  $h$ , and  $d$  are obtained from previous statistics.

We formulate the three models above as follows. Sets of the form  $C = \{c_1, \dots, c_k\}$  used in this work are equipped with linear order  $c_i < c_{i+1}$  for  $1 \leq i < k$ . We index the population of  $n$  items from 1 to  $n$ , namely  $\mathcal{N} = \{1, 2, \dots, n\}$ . Let  $\mathbf{x} = (x_1, \dots, x_n)^T \in \{-1, 0, 1\}^n$  be the binary representation vector of  $n$  items, where  $x_j = 1$  indicates that item  $j$  is positive,  $x_j = 0$  indicates that item  $j$  is negative, and  $x_j = -1$  indicates that item  $j$  is inhibitory. A test on a subset of items is positive if the subset contains at least one positive and does not contain any inhibitors. Otherwise, the test outcome is negative.

The test notation is denoted as  $\odot$ . Let  $\mathbf{p} = (p_1, \dots, p_n) \in \{0, 1\}^n$  be the test representation vector. Then, the outcome vector of the test  $\mathbf{p}$  with the input vector  $\mathbf{x}$ , namely  $\mathbf{p} \odot \mathbf{x}$ , is positive (1) if there does not exist a  $j$  such that  $p_j = 1$  and  $x_j = -1$ , and there

exists a  $j'$  such that  $p_{j'} = 1$  and  $x_{j'} = 1$ . The test outcome is negative (0) otherwise. Given a measurement matrix  $\mathcal{M}$  of size  $t \times n$  and an input vector  $\mathbf{x}$ , the corresponding outcome vector is  $\mathcal{M} \odot \mathbf{x} = [y_1, \dots, y_n]^T$ , where  $y_i = \mathcal{M}(i, :) \odot \mathbf{x}$ .

There are two common decoding types based on classification strategy. The first is to only identify the positives while the second is to identify both positives and inhibitors. Our objective is to find an efficient encoding and decoding scheme to satisfy two decoding types, i.e., minimizing the number of tests and the decoding time.



**Figure 1.** Three models for blocks of positives and inhibitors. Red, purple, and black dots represent positives, inhibitors, and negative items, respectively. A double arrow line stands for a block of  $D$  consecutive items. The first, second, and third models are a single block of consecutive positives and inhibitors, single block of positives and inhibitors, and blocks of positives and inhibitors. The second model is a generalization of the first model, and the third model is a generalization of the first two models.

### 1.2. Contributions

**Overview:** We study group testing with blocks of positives and inhibitors and provide efficient encoding and decoding schemes to tackle it. By leveraging the knowledge of positives and inhibitors belonging to a small interval of size  $D$ , our objective is to identify the position of some positive, say  $j^*$ ; then, one could claim that the indices of all positives and inhibitors must belong to the range from  $\max\{1, j^* - D + 1\}$  to  $\min\{j^* + D - 1, n\}$ . To precisely identify positives and inhibitors, appropriate tests are designed to accomplish this task.

Our proposed scheme includes two procedures, which are the filtering and scrutinizing procedures. The tests in the filtering procedure remove most negative items and leave a subset(s) of size up to  $2D$  that contains all positives, inhibitors, and probably some negatives. The tests in the scrutinizing procedure remove all negatives and then identify positives and inhibitors. The details of the two procedures are specified in accordance to each specific problem.

The contributions for a single block of (consecutive) positives and inhibitors is summarized in Theorem 1. The proofs for the results of the first and second model are described in Sections 3 and 4.

**Theorem 1.** Let  $1 \leq d, h, d + h \leq n$  be integers. Suppose that a population of  $n$  linearly ordered items includes exactly  $d$  positives and  $h$  inhibitors in a block of  $D \geq d + h$  items that are consecutive that order. When  $D = d + h$  (respectively,  $D \geq d + h$ ), there exists a deterministic and strongly explicit measurement matrix of size  $O\left(h \log \frac{hn}{d+h} + d\right) \times n$  (respectively,  $O(D \log n) \times n$ ) that can be used to identify all positives in  $O\left(h \log \frac{hn}{d+h} + d\right)$  (respectively,  $O(D \log n)$ ) time. Moreover, it requires  $O\left((d + h)^3 \log \frac{n}{d+h}\right)$  (respectively,  $O(D \log n + D^3 \log(n/D))$ ) tests to identify all positives and inhibitors in  $O\left((d + h)^4 \log \frac{n}{d+h}\right)$  (respectively,  $O(D \log n + D^4 \log(n/D))$ ) time.

The contribution for blocks of positives and inhibitors is summarized in the following theorem, which is proved later in Section 5.

**Theorem 2.** Let  $1 \leq d, h, d + h \leq D \leq n$  be integers. Suppose that a population of  $n$  items is linearly ordered and the positives and inhibitors belong to blocks of consecutive items in which each block has a size of up to  $D$  and contains up to  $d$  positives and  $h$  inhibitors. Then, there exists a deterministic and strongly explicit measurement matrix of size  $O(Dk^2(D + \log \frac{n}{D}) \log \frac{n}{D}) \times n$  that can be used to identify all positives in  $O(Dk^2(D + \log \frac{n}{D}) \log \frac{n}{D})$  time. Moreover, it requires  $O(D^2k^2(D + \log \frac{n}{D}) \log \frac{n}{D})$  tests to identify all positives and inhibitors in time.

$$O\left(Dk^2 \log \frac{n}{kD} (D + \log \frac{n}{D}) + k^4 D^4 \log \frac{n}{kD}\right).$$

### 2. Preliminaries

Disjunct matrices were first introduced by Kautz and Singleton [28] as *superimposed codes* and then generalized by Stinson and Wei [29] and D’yachkov et al. [30]. We later use them for identifying both positives and inhibitors. Let the support set for vector  $\mathbf{v} = (v_1, \dots, v_w)$  be  $\text{supp}(\mathbf{v}) = \{j \mid v_j \neq 0\}$  and  $|\mathbf{v}| = |\{j \mid v_j \neq 0\}|$ . We denote  $\mathcal{M}(i, \cdot)$  and  $\mathcal{M}(\cdot, j)$  as the  $i$ th row and the  $j$ th column of matrix  $\mathcal{M}$ . The formal definition of a disjunct matrix is as follows.

**Definition 1.** An  $m \times n$  binary matrix  $\mathcal{M}$  is called an  $(n, v, u)$ -disjunct matrix if, for any two disjoint subsets  $S_1, S_2 \subset [n]$  such that  $|S_1| = v$  and  $|S_2| = u$ , there exists at least one row in which there are all 1s among the columns in  $S_2$  while all the columns in  $S_1$  have 0s, i.e.,  $|\bigcap_{j \in S_2} \text{supp}(\mathcal{M}(\cdot, j)) \setminus \bigcup_{j \in S_1} \text{supp}(\mathcal{M}(\cdot, j))| \geq 1$ .

Chen et al. [31] gave an upper bound on the number of rows for  $(n, v, u)$ -disjunct matrices as follows.

**Theorem 3** ([31] Theorem 3.2). For any positive integers  $v, u$ , and  $n$  with  $x = v + u \leq n$ , there exists a  $t \times n$   $(n, v, u)$ -disjunct matrix with the following.

$$t(n, v, u) = O\left(\left(\frac{x}{u}\right)^u \left(\frac{x}{v}\right)^v x \log \frac{n}{x}\right). \tag{1}$$

Once  $u = 1$ ,  $(n, v, 1)$ -disjunct matrices become  $v$ -disjunct matrices. The following theorem states the construction and decoding time for a  $d$ -disjunct matrix.

**Theorem 4** ([11] Theorem 16). Let  $1 \leq d \leq n$ . Then, there exists a deterministic and explicit  $t \times n$   $d$ -disjunct matrix with  $t = O(d^2 \log n)$  that can be decoded in the polynomial time of  $t$ .

### 3. Single Block of Consecutive Positives and Inhibitors

In this section, we consider the case when the positives and inhibitors are consecutive and the numbers of positives and inhibitors are known in advance. Set  $D = d + h$ .

#### 3.1. Encoding Procedure

Set  $a = \lfloor D/(h + 1) \rfloor$  and  $\kappa = \lceil n/a \rceil$ . A super item, denoted as  $\bar{j}$ , is a set of consecutive items. The  $n$  items are distributed into  $\kappa$  super items indexed from 1 to  $\kappa$  and each super item contains exactly  $a$  items, except that the last one may contain less than  $a$  items. Let  $\bar{\mathcal{E}} = \{\bar{1}, \bar{2}, \dots, \bar{\kappa}\}$  be the set of super items generated from  $\mathcal{N}$ , where  $\text{set}(\bar{j}) = \{(j - 1)a + 1, \dots, ja\}$  for  $j = 1, \dots, \kappa - 1$  and  $\text{set}(\bar{\kappa}) = \{(\kappa - 1)a + 1, \dots, n\}$ . We then denote that  $\chi_{\bar{\mathcal{E}}} = (\chi_1, \dots, \chi_\kappa)$  be the characteristic vector of  $\bar{\mathcal{E}}$ , where  $\chi_j = 1$  if the test on  $\bar{j}$  is positive and  $\chi_j = 0$  otherwise.

##### 3.1.1. Filtering Matrices

We create  $h + 2$  filtering matrices in the filtering procedure as follows. Let  $\mathcal{F} = [\mathbf{f}_1, \dots, \mathbf{f}_\kappa]$  be an  $f \times \kappa$  (indexing) binary matrix for which its  $j$ th column is the  $f$ -bit binary representation of integer  $j$ , where  $f = \lceil \log(\kappa + 1) \rceil$ . It is obvious that the index  $j$  is

uniquely identified by  $\mathbf{f}_j$ . We then generate  $h + 2$  binary matrices  $\mathcal{F}^{(u)} = [\mathcal{F}_1^{(u)}, \dots, \mathcal{F}_\kappa^{(u)}]$  for  $u = 1, \dots, h + 2$ , such that column  $\mathcal{F}^{(u)}(:, j)$  is a zero vector if  $j \not\equiv u \pmod{h + 2}$  while  $\mathcal{F}^{(u)}(:, j) = \mathbf{f}_j$  if  $j \equiv u \pmod{h + 2}$ . For example, let  $n = 12, d = 4$ , and  $h = 2$ . We obtain  $a = \lceil (d + h) / (h + 1) \rceil = 2$  and  $\kappa = \lceil n/a \rceil = 6$ . Since  $\mathcal{F} = [\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4, \mathbf{f}_5, \mathbf{f}_6]$ , we imply  $\mathcal{F}^{(1)} = [\mathbf{f}_1, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{f}_5, \mathbf{0}]$ ,  $\mathcal{F}^{(2)} = [\mathbf{0}, \mathbf{f}_2, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{f}_6]$ ,  $\mathcal{F}^{(3)} = [\mathbf{0}, \mathbf{0}, \mathbf{f}_3, \mathbf{0}, \mathbf{0}, \mathbf{0}]$ , and  $\mathcal{F}^{(4)} = [\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{f}_4, \mathbf{0}, \mathbf{0}]$ . For every  $h + 2$  consecutive column in  $\mathcal{F}^{(u)}$ , there exists only one non-zero column. Therefore, it is used to “isolate” each super item in the  $h + 1$  super items generated from the set of  $D$  positives and inhibitors.

Let  $\mathbf{y}_{\mathcal{F}^{(u)}} = [y_{\mathcal{F}^{(u)}}(1), \dots, y_{\mathcal{F}^{(u)}}(f)]^T$  be the outcome vector by using the testing matrix  $\mathcal{F}^{(u)}$  with the set of super items  $\bar{\mathcal{E}}$ . In particular, if  $\mathcal{F}^{(u)}(i, j) = 1$  (respectively,  $\mathcal{F}^{(u)}(i, j) = 0$ ) then all items in the super item  $\bar{j}$  (respectively, do not) belong to test  $i$ . Therefore, we obtain the following.

$$\mathbf{y}_{\mathcal{F}^{(u)}} = \mathcal{F}^{(u)} \odot \chi_{\bar{\mathcal{E}}}. \tag{2}$$

### 3.1.2. Sanitizing Matrices

In the sanitizing procedure, the measurement matrix depends on whether the objective is to identify positives only or to identify both positives and inhibitors. For the first objective, we design an  $s \times n$  matrix  $\mathcal{S}$  such that  $\mathcal{S}(i, j) = 1$  if  $i \equiv j \pmod{s}$  and  $\mathcal{S}(i, j) = 0$  or, otherwise, where  $s = 2D - 1$ . In other words, each test contains items spaced  $2D - 1$  apart in a linear order. For example, when  $n = 12, d = 4$ , and  $h = 2$ , we obtain the following.

$$\mathcal{S} = \left[ \begin{array}{cccccccccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array} \right].$$

It is straightforward that every column in  $\mathcal{S}$  (respectively,  $\mathcal{F}^{(u)}$ ) is deterministic and strongly explicit because each column in it can be generated in time and space of  $O(2D) = O(d + h)$  (respectively,  $O(\log \kappa)$ ).

For the second objective, i.e., the objective of identifying both positives and inhibitors, we design an additional matrix  $\mathcal{R}$  along with matrix  $\mathcal{S}$ . Let  $\mathcal{R}$  be a  $r \times n$  ( $n, 2D - 3, 2$ )-disjunct matrix as defined in Definition 1. Therefore, we have  $r = O(D^3 \log(n/D)) = O((d + h)^3 \log(n/(d + h)))$  as in Theorem 3.

Let  $\mathbf{y}_{\mathcal{S}} = [y_{\mathcal{S}}(1), \dots, y_{\mathcal{S}}(s)]^T$  (respectively,  $\mathbf{y}_{\mathcal{R}} = [y_{\mathcal{R}}(1), \dots, y_{\mathcal{R}}(r)]^T$ ) be the outcome vector by using the testing matrix  $\mathcal{S}$  (respectively,  $\mathcal{R}$ ) with input set  $N$ . In particular, we have the following.

$$\mathbf{y}_{\mathcal{S}} = \mathcal{S} \odot \mathbf{x} \text{ and } \mathbf{y}_{\mathcal{R}} = \mathcal{R} \odot \mathbf{x}. \tag{3}$$

### 3.2. Decoding Procedure and Correctness

We first approximately locate some positive items by using outcome vectors  $\mathbf{y}_{\mathcal{F}^{(1)}}, \dots, \mathbf{y}_{\mathcal{F}^{(h+1)}}$ . Then, we can locate a set of up to  $2D - 1$  items that contains all positives and inhibitors. We call this set *the set of interest*. By using  $\mathbf{y}_{\mathcal{S}}$ , we can exactly identify all positives in that set. Meanwhile, if  $\mathbf{y}_{\mathcal{R}}$  is also used, all inhibitors are also identified. The details of the decoding procedure are as follows.

Let  $\lambda$  be an index such that  $\mathbf{y}_{\mathcal{F}^{(\lambda)}}$  is not a zero vector. There always exists a  $\lambda$ . Indeed, because there are  $h$  inhibitors,  $D$  consecutive positives and inhibitors, and each super item contains up to  $\lfloor D/(h + 1) \rfloor$  items, the total number of items contained in super items

having inhibitors is up to  $h \lfloor D/(h+1) \rfloor < D$ . Therefore, there must exist a super item  $\bar{\alpha}$  that does not contain any inhibitor but all positives for  $1 \leq \alpha \leq \kappa$ . Let  $\lambda$  be the index such that  $\mathcal{F}^{(\lambda)}(:, \alpha) \neq \mathbf{0}$ . Since two consecutive non-zero column in  $\mathcal{F}^{(u)}$  are space by  $h+2$ ,  $\mathbf{y}_{\mathcal{F}^{(\lambda)}} = \mathcal{F}^{(\lambda)} \odot \chi_{\mathcal{E}} = \mathcal{F}^{(\lambda)}(:, \alpha)$ . Therefore, to identify  $\alpha$ , we convert a non-zero vector  $\mathbf{y}_{\mathcal{F}^{(\lambda)}}$  into a decimal number. The indices of all positives and inhibitors, thus, must belong to the range from  $\max\{1, j^* - D + 1\}$  to  $\min\{j^* + D - 1, n\}$ . The decoding complexity of identifying  $\lambda$  is therefore  $O(hf)$ .

Because of the construction of  $\mathcal{S}$ , a matrix composed of  $2D - 1$  consecutive columns in it is a permutation of a  $(2D - 1) \times (2D - 1)$  identity matrix. Therefore, given the indices from  $\max\{1, \alpha - D + 1\}$  to  $\min\{\alpha + D - 1, n\}$ , one can identify which item is positive based on the corresponding outcome vector  $\mathbf{y}_{\mathcal{S}}$ . The decoding complexity of  $\mathbf{y}_{\mathcal{S}}$  is, therefore,  $O(D) = O(d + h)$ .

After identifying  $d$  positives, the set of interest contains up to  $2D - 1 - d = d + 2h - 1$  potential inhibitors. Because of the construction of  $\mathcal{R}$ , for any two items and other  $2D - 3$  items, there exists a test that contains the two items and does not contain the other  $2D - 3$  items. Therefore, one could identify whether a potential inhibitor is truly an inhibitor by checking the row that contains it and whether it is a positive, in addition to checking that it does not contain the remaining items in the set of interest. Since there are up to  $2D - 1 - d$  potential positives and the number of rows in  $\mathcal{R}$  is  $r$ , this procedure to identify inhibitors takes  $O(r(2D - 1 - d)) = O(r(d + h))$ .

### 3.3. Decoding Complexity and Number of Tests

As analyzed in the previous section, to identify the positives only, the number of required tests and the decoding complexity are as follows.

$$O((h+2)f+s) = O\left(h \log \frac{hn}{d+h}\right) + O(d+h) = O\left(h \log \frac{hn}{d+h} + d\right).$$

To identify both the positives and inhibitors, i.e., classify all items, the required number of tests is as follows.

$$\begin{aligned} (h+2)f + s + r &= O\left(h \log \frac{hn}{d+h}\right) + O(d+h) + O\left((d+h)^3 \log \frac{n}{d+h}\right) \\ &= O\left((d+h)^3 \log \frac{n}{d+h}\right). \end{aligned}$$

The corresponding decoding complexity is as follows.

$$O(hf) + O(d+h) + O(r(d+h)) = O\left((d+h)^4 \log \frac{n}{d+h}\right).$$

## 4. Single Block of Positives and Inhibitors

In this section, we consider the case when the positives and inhibitors are not necessarily consecutive but belong to a small block (set) of consecutive items of size up to  $D \geq d + h$ , where  $d$  and  $h$  are the maximum numbers of positives and inhibitors in the population of  $n$  items.

In the encoding procedure, we use the same techniques in Section 3.1 but adjust some parameters. In the filtering procedure, we set  $a = 1$ , i.e., every super item reduces to an item. Therefore,  $\kappa$  is equal to  $n$ . Moreover, we create  $D$  filtering matrices, i.e.,  $h + 1$  is replaced by  $D$ . In the sanitizing procedure, the parameter  $s$  in the  $s \times n$  matrix  $\mathcal{S}$  is set to be  $2D - 1$ . Matrix  $\mathcal{R}$  is a  $r \times n$   $(n, 2D - 3, 2)$ -disjunct matrix as defined in Definition 1. Therefore, we have  $r = O(D^3 \log(n/D))$  as in Theorem 3.

Since the decoding procedure and the proofs of correctness are as the same as in Section 3.2, we only pay attention for the required numbers of tests and the decoding complexities. Each matrix  $\mathcal{F}^{(u)}$  has a size of  $f \times n$ , where  $f = \lceil \log n \rceil$ , for  $u = 1, \dots, D$ .

The numbers of tests in matrices  $\mathcal{S}$  and  $\mathcal{R}$  are  $s = 2D - 1$  and  $r = O(D^3 \log(n/D))$ , respectively.

To identify the positives only, the number of required tests and the decoding complexity are as follows.

$$Df + s = O(D \log n) + O(D) = O(D \log n).$$

To identify both the positives and inhibitors, the required number of tests is described as follows.

$$\begin{aligned} Df + s + r &= O(D \log n) + O(D) + O(D^3 \log(n/D)) \\ &= O(D \log n + D^3 \log(n/D)). \end{aligned}$$

The corresponding decoding complexity is as follows.

$$O(Df) + O(d + h) + O(rD) = O(D \log n + D^4 \log(n/D)).$$

### 5. Blocks of Positives and Inhibitors

In this section, we consider a model consisting of multiple blocks of positives and inhibitors, in which all positives and inhibitors belong to at most  $k$  special blocks of consecutive items and each block has up to  $D$  consecutive items. Moreover, each special block contains up to  $d$  positives and up to  $h$  inhibitors.

#### 5.1. Encoding Procedure

We generate  $D$  sets from the set of  $n$  items  $\mathcal{N}$  as follows. Set  $N^{(u)} = \{u, D + u, 2D + u, \dots, n_u\}$  and  $\mathbf{x}^{(u)} = (x_u, x_{D+u}, x_{2D+u}, \dots, x_{n_u})^T$ , where  $n_u$  is the largest number smaller than  $n$  and  $n_u \equiv u \pmod{D}$ , for  $u = 1, \dots, D$ . It is obvious that  $n_u = |N^{(u)}| \leq \lceil n/D \rceil$ . Since each special block has up to  $D$  items and there are up to  $k$  special blocks, each set  $N^{(u)}$  must contain up to  $k$  positives and inhibitors in total. Moreover, for each special block  $\tau$ , there exists an index  $u_\tau$  such that some positive item in that block belongs to  $N^{(u_\tau)}$  because two consecutive items in  $N^{(u)}$  are spaced apart by  $D$  and each special block has up to  $D$  items.

Let  $\mathcal{M}^{(u)}$  be an  $m_u \times n_u$   $k$ -disjunct matrix. We then obtained  $m_u = O(k^2 \log n_u) = O(k^2 \log(n/D))$  as in Theorem 4. Let  $\mathcal{B}^{(u)}$  be a  $b \times n_u$  index matrix:

$$\mathcal{B}^{(u)} := \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_{n_u} \\ \bar{\mathbf{b}}_1 & \bar{\mathbf{b}}_2 & \dots & \bar{\mathbf{b}}_{n_u} \end{bmatrix} = [\mathcal{B}_1^{(u)} \quad \dots \quad \mathcal{B}_{n_u}^{(u)}], \tag{4}$$

where  $b = 2\lceil \log n_u \rceil$ ,  $\mathbf{b}_j$  is the  $\lceil \log n_u \rceil$ -bit binary representation of integer  $j - 1$ ,  $\bar{\mathbf{b}}_j$  is the complement of  $\mathbf{b}_j$ , and  $\mathcal{B}_j^{(u)} := \begin{bmatrix} \mathbf{b}_j \\ \bar{\mathbf{b}}_j \end{bmatrix}$  for  $j = 1, 2, \dots, n_u$ . Item  $j$  is characterized by column  $\mathcal{B}_j$  and that the weight of every column in  $\mathcal{B}$  is  $b/2 = \lceil \log n_u \rceil$ . Furthermore, the index  $j$  is uniquely identified by  $\mathbf{b}_j$ . For example, if we set  $n_u = 8$ ,  $b = 2\lceil \log n_u \rceil = 6$ , and the matrix in (4) becomes the following.

$$\mathcal{B}^{(u)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}. \tag{5}$$

Finally, matrices  $\mathcal{S}$  and  $\mathcal{R}$  are defined as in Section 3.1.2. Note that  $D$  is not set to be  $d + h$  here.

We are now ready to generate a filtering matrix and a scrutinizing matrix. The filtering matrix corresponding to matrix  $\mathcal{M}^{(u)}$  is as follows:

$$\mathcal{F}^{(u)} = \begin{bmatrix} \mathcal{M}^{(u)}(1,:) \\ \mathcal{B}^{(u)} \times \text{diag}(\mathcal{M}^{(u)}(1,:)) \\ \vdots \\ \mathcal{M}^{(u)}(m_u,:) \\ \mathcal{B}^{(u)} \times \text{diag}(\mathcal{M}^{(u)}(m_u,:)) \end{bmatrix}, \tag{6}$$

where  $\text{diag}(\cdot)$  is a diagonal matrix generated by the input vector.

The vector observed after performing the tests given by the measurement matrix  $\mathcal{F}^{(u)}$  is described as follows:

$$\mathbf{y}^{(u)} = \mathcal{F}^{(u)} \odot \mathbf{x}^{(u)} = \begin{bmatrix} \mathcal{M}^{(u)}(1,:) \odot \mathbf{x}^{(u)} \\ \mathcal{B}^{(u)} \odot \mathbf{x}_1^{(u)} \\ \vdots \\ \mathcal{M}^{(u)}(m_u,:) \odot \mathbf{x}^{(u)} \\ \mathcal{B}^{(u)} \odot \mathbf{x}_{m_u}^{(u)} \end{bmatrix} = \begin{bmatrix} y_1^{(u)} \\ \mathbf{y}_1^{(u)} \\ \vdots \\ y_{m_u}^{(u)} \\ \mathbf{y}_{m_u}^{(u)} \end{bmatrix} \tag{7}$$

where  $\mathbf{x}_i^{(u)} = \text{diag}(\mathcal{M}^{(u)}(i,:)) \times \mathbf{x}^{(u)}$ ,  $y_i = \mathcal{M}^{(u)}(i,:) \odot \mathbf{x}^{(u)}$ , and  $\mathbf{y}_i^{(u)} = \mathcal{B}^{(u)} \odot \mathbf{x}_i^{(u)}$ , for  $i = 1, 2, \dots, m_u$ . Entry  $y_i$  indicates whether there exist only negatives and positives in that test. If the answer is yes, vector  $\mathbf{y}_i^{(u)}$  tells us whether there exists only one positive or more than one positive.

Let  $\text{expand}(\mathcal{M}^{(u)}(i,:))$  be  $\mathcal{M}^{(u)}(i,:)$ . Then, for any  $j \in N^{(u)}$  and  $\mathcal{M}^{(u)}(i,j) = 1$ , every entry in  $\text{expand}(\mathcal{M}^{(u)}(i,:))$  indexed from  $\max\{j - D + 1, 1\}$  to  $\min\{j + D - 1, n\}$  is set to be 1. This vector is used to identify a block of  $2D - 1$  consecutive items that contains at least one positive item. In particular, to identify positives only, the scrutinizing matrix corresponding to matrix  $\mathcal{M}^{(u)}$  is defined as follows:

$$\mathcal{S}^{(u)} = \begin{bmatrix} \mathcal{S} \times \text{diag}(\text{expand}(\mathcal{M}^{(u)}(1,:))) \\ \vdots \\ \mathcal{S} \times \text{diag}(\text{expand}(\mathcal{M}^{(u)}(m_u,:))) \end{bmatrix}, \tag{8}$$

where  $\mathcal{S}$  is defined in Section 3.1.2, and the outcome vector obtained by using this matrix is as follows:

$$\mathbf{s}^{(u)} = \mathcal{S}^{(u)} \odot \mathbf{x} = \begin{bmatrix} \mathcal{S} \odot (\text{diag}(\text{expand}(\mathcal{M}^{(u)}(1,:))) \times \mathbf{x}) \\ \vdots \\ \mathcal{S} \odot (\text{diag}(\text{expand}(\mathcal{M}^{(u)}(m_u,:))) \times \mathbf{x}) \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1^{(u)} \\ \vdots \\ \mathbf{s}_{m_u}^{(u)} \end{bmatrix}, \tag{9}$$

where  $\mathbf{s}_i^{(u)} = \mathcal{S} \odot (\text{diag}(\text{expand}(\mathcal{M}^{(u)}(i,:))) \times \mathbf{x})$ , for  $i = 1, 2, \dots, m_u$ .

To identify both positives and inhibitors, an additional scrutinizing  $(n, kD - 2, 2)$ -disjunct matrix  $\mathcal{R}$  is used. Let  $\mathbf{r}$  be the outcome vector by using this matrix.

### 5.2. Decoding Procedure and Correctness

For each  $u \in \{1, D\}$ , we first scan each  $\mathbf{y}^{(u)}$  to locate some positive item in some block. The decoding procedure is as follows. First, find  $1 \leq i \leq m_u$  such that  $y_i^{(u)} = 1$  and  $|\mathbf{y}_i^{(u)}| = \lceil \log n_u \rceil$ . Second, let  $\alpha$  be the corresponding decimal number of the first half of  $\mathbf{y}_i^{(u)}$ , where  $y_i^{(u)} = 1$  and  $|\mathbf{y}_i^{(u)}| = \lceil \log n_u \rceil$ . Then, similarly to the arguments in Section 3.2, since any matrix composed of  $2D - 1$  consecutive columns in  $\mathcal{S}$  is a permutation of a  $(2D - 1) \times (2D - 1)$  identity matrix, one can identify which item is positive based on the

corresponding outcome vector  $\mathbf{s}_i^{(u)}$ . Finally, all inhibitors in a block can be identified by using  $\mathbf{r}_i^{(u)}$ .

Such  $i$  always exists in the first step. Indeed, as proved in Section 5.1 that for each special block  $\tau$ , there exists an index  $u_\tau$  such that some positive item in that block belongs to  $N^{(u_\tau)}$ . Since each set  $N^{(u_\tau)}$  contains up to  $k$  positives and inhibitors and  $\mathcal{M}^{(u_\tau)}$  is a  $k$ -disjunct matrix, there must exist row  $i$  such that  $\mathcal{M}^{(u_\tau)}(i, :)$  contains only that positive. Therefore,  $y_i^{(u_\tau)} = 1$  and  $|\mathbf{y}_i^{(u_\tau)}| = \lceil \log n_{u_\tau} \rceil$ . Conversely, if  $y_i^{(u)} = 1$ , there must exist at least one positive item in  $N^{(u)}$  in test  $\mathcal{M}^{(u)}(i, :)$ . Moreover, since  $\mathcal{B}^{(u)} \odot (\text{diag}(\mathcal{M}^{(u)}(i, :)) \times \mathbf{x}^{(u)})$  and every column in  $\mathcal{B}^{(u)}$  has weight of  $\lceil \log n_u \rceil$ , there must exist only one positive item in  $N^{(u)}$  in that test. Otherwise,  $|\mathbf{y}_i^{(u)}| > \lceil \log n_u \rceil$ .

In the second step, the indices of positives and inhibitors then ranged from  $\max\{1, \alpha - D + 1\}$  to  $\min\{\alpha + D - 1, n\}$ . Because of the construction of vector  $\text{expand}(\mathcal{M}^{(u)}(1, :))$ , every item indexed from  $\max\{1, \alpha - D + 1\}$  to  $\min\{\alpha + D - 1, n\}$  presents in the characteristic vector  $\text{diag}(\text{expand}(\mathcal{M}^{(u)}(1, :))) \times \mathbf{x}$ . Therefore,  $\mathbf{s}_i^{(u)}$  is the union of up to  $D$  columns in  $\mathcal{S}$ , which out of them corresponds to all positives and inhibitors in a specific block. The positives are thus identified. The decoding complexity of  $\mathbf{s}_i^{(u)}$  is therefore  $O(D)$ .

In the last step, since  $\mathcal{R}$  is an  $(n, kD - 2, 2)$ -disjunct matrix, for any block of positives and inhibitors, there exists a row such that it contains only a positive and an inhibitor. That inhibitor is thus identified. This procedure takes  $O(k \times r(2D - 1)) = O(k^4 D^4 \log(n / (kD)))$ .

### 5.3. Decoding Complexity and Number of Tests

There are  $D \mathcal{F}^{(u)}$  deterministic and strongly explicit matrices in the filtering procedure. Since each has  $m_u(1 + b) = O(k^2 \log n_u \times \log n_u) = O(k^2 \log^2(n/D))$  tests, the total number of tests in the filtering procedure is  $O(Dk^2 \log^2(n/D))$ . The decoding complexity by using these tests is also  $O(Dk^2 \log^2(n/D))$ .

There are also  $D \mathcal{S}^{(u)}$  matrices and  $D \mathcal{R}^{(u)}$  matrices. The total number of tests for  $D \mathcal{S}^{(u)}$  matrices and  $D \mathcal{R}^{(u)}$  matrices include  $m_u D(2D - 1) = O(m_u D^2)$  and  $m_u D r = O(m_u r D)$ , respectively. Therefore, the number of tests for identifying positives only (both positives and inhibitors) is  $O(Dm_u(1 + b + s)) = O(Dk^2 \log(n/D)(D + \log(n/D)))$  (respectively,  $O(Dm_u(1 + b + s) + r) = O(k^2 D^2 \log(n/D)(\log(n/D) + kD))$ ).

For each  $u$ , the running time to decode all  $\mathbf{s}_i^{(u)}$ s is  $O(m_u s) = O(Dk^2 \log(n/D))$ . Since  $u$  ranges from 1 to  $D$ , the running time to find all positives is  $O(Dk^2 \log^2(n/D)) + D \times O(m_u s) = O(Dk^2 \log(n/D)(D + \log(n/D)))$ . On the other hand, the running time to find all positives and inhibitors is as follows.

$$O\left(Dk^2 \log \frac{n}{kD} (D + \log \frac{n}{D})\right) + O\left(k^4 D^4 \log \frac{n}{kD}\right) = O\left(Dk^2 \log \frac{n}{kD} (D + \log \frac{n}{D}) + k^4 D^4 \log \frac{n}{kD}\right).$$

## 6. Discussion

### 6.1. Comparison

We compare our proposed schemes with existing schemes, namely Ganesan et al. [32], Chang et al. [33], and Bui et al. [20] in Table 1. There are eight criteria to consider here. The first four criteria are about the structure of the population of  $n$  items. They are the number of blocks, the number of items in a block, the number of positives (in a block if applicable), and the number of inhibitors (in a block if applicable). The fifth criterion is the decoding type. The sixth is the construction type, which describes how measurement matrices can be achieved. The seventh and the last are the number of tests and the decoding time.

Consider the decoding type as “positives only.” The construction type in our proposed schemes for the first and the second model, i.e., the number of blocks is one, is deterministic and strongly explicit. They are better than the schemes proposed by Chang et al. and Gane-

san et al., whose schemes are random and explicit. The numbers of tests in our proposed schemes are almost less than a factor of  $d + h$  compared to the ones in Chang et al.’s and Bui et al.’s schemes and less than the one in Ganesan et al.’s scheme. More importantly, our decoding times scale to the number of tests while the ones in Chang et al.’s and Ganesan et al.’s schemes scale to the number of tests and the number of items. For the third model, the same arguments are applied by replacing  $d$  by  $dk$  and  $h$  by  $hk$ .

We now consider the decoding type as “positives and inhibitors.” For the first and second models, the number of tests in our proposed schemes is relatively the same as the one in Chang et al.’s scheme, smaller than the one in Bui et al.’s scheme, and less than the one in Ganesan et al.’s scheme. Meanwhile, the decoding times are smaller than the ones in the three existing schemes. For the third model, the number of tests in our proposed schemes is relatively similar to the one in Chang et al.’s scheme, smaller than the one in Bui et al.’s scheme, and larger than the one in Ganesan et al.’s scheme. However, our decoding time is smaller than the ones in the three existing works.

**Table 1.** Comparison with previous work. “Det.” and “Rnd.” stand for “Deterministic” and “Random.” We set  $\lambda = \frac{(d+h) \ln n}{W((d+h) \log n)}$ ,  $\alpha = \max\left\{\frac{\lambda}{(d+h)^2}, 1\right\}$ , and  $\beta = O(Dk^2(D + \log \frac{n}{D}) \log \frac{n}{D} + k^4 D^4 \log \frac{n}{kD})$ , where  $W(x)e^{W(x)} = x$  and  $W(x) \sim \Theta(\log x - \log \log x)$ .

No. of Blocks	No. of Items in a Block	No. of Positives (in a Block)	No. of Inhibitors (in a Block)	Decoding Type	Scheme	Construction Type	No. of Tests $t$	Decoding Complexity
Not applicable		$d$	$h$	Positives only	Ganesan et al. [32]	Rnd., Explicit	$O((d + h) \log n)$	$O(tn)$
		$\leq d$	$\leq h$		Chang et al. [33]		$O((d + h)^2 \log n)$	$O(tn)$
		$\leq d$	$\leq h$		Bui et al. [20]		$O(\lambda^2 \log n)$	$O\left(\frac{\lambda^5}{(d+h)^2}\right)$
1	$d + h$	$d$	$h$	Positives only	<b>Theorem 1</b>	Det., Strongly explicit	$O\left(h \log \frac{hn}{d+h} + d\right)$	$O(t)$
	$D$	$\leq d$	$\leq h$		<b>Theorem 1</b>		$O(D \log n)$	$O(t)$
$k$	$\leq D$	$\leq d$	$\leq h$	Positives only	<b>Theorem 2</b>	Rnd., Explicit	$O(Dk^2(D + \log \frac{n}{D}) \log \frac{n}{D})$	$O(t)$
Not applicable		$d$	$d$	Positives and inhibitors	Ganesan et al. [32]	Rnd., Explicit	$O((d + h^2) \log n)$	$O(tn)$
		$\leq d$	$\leq h$		Chang et al. [33]		$O((d + h)^3 \log n)$	$O(tn)$
		$\leq d$	$\leq h$		Bui et al. [20]		Det., Strongly explicit	$O(\lambda^3 \log n)$
1	$d + h$	$d$	$h$	Positives and inhibitors	<b>Theorem 1</b>	Rnd., Explicit	$O\left((d + h)^3 \log \frac{hn}{d+h}\right)$	$O\left((d + h)^4 \log \frac{hn}{d+h}\right)$
	$D$	$\leq d$	$\leq h$		<b>Theorem 1</b>		$O(D \log n + D^3 \log \frac{n}{D})$	$O(D \log n + D^4 \log \frac{n}{D})$
$k$	$\leq D$	$\leq d$	$\leq h$	Positives and inhibitors	<b>Theorem 2</b>	Rnd., Explicit	$O(D^2 k^2(D + \log \frac{n}{D}) \log \frac{n}{D})$	$\beta$

### 6.2. Potential Applications

Bruno et al. [34] addressed a group testing-based solution in genetic mapping and sequencing. In this application, the authors consider linear DNA, which consists of consecutive segments of the DNA. Each segment is placed in a pool, called clones, in an order consistently to the order of their appearance in the linear DNA. A collection of such clones is called a linear DNA library. From this point, we can ask where segments (clones) of interest are in the linear DNA library [35]. The segments of interest here can be considered as positives and other segments can be considered as negatives. A pool that contains at least one segment of interest returns a positive outcome when performing testing and returns a negative outcome otherwise.

We extend the application above to a potential application as follows. Given a linear DNA library, we would like to find segments of DNA that express a certain biological property and segments of DNA that inhibit the segments expressing a certain biological property. The first and second types of segments of interest are considered as positives and inhibitors, respectively, while the remaining segments are considered as negatives. Because of the nature of DNA, an inhibitor is usually close to positives. Therefore, the blocks of positives and inhibitors model can be used to identify both positives and inhibitors.

## 7. Conclusions

In this paper, we presented efficient encoding and decoding procedures to identify positives and/or inhibitors in a single block of (consecutive) positives and inhibitors or in blocks of positives and inhibitors. The number of tests and the decoding times in our proposed schemes is usually smaller than the ones in existing works. An extension of this work to other settings in group testing such as threshold group testing or complex group testing is still an open problem.

**Author Contributions:** Conceptualization, T.V.B.; methodology, T.V.B.; writing—original draft preparation, T.V.B.; writing—review and editing, T.V.B., I.E., M.K., T.K., and T.D.N.; funding acquisition, T.D.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** Thuc D. Nguyen was funded in part by University of Science, VNU-HCM under Grant No. CNTT 2021-27.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Dorfman, R. The detection of defective members of large populations. *Ann. Math. Stat.* **1943**, *14*, 436–440. [\[CrossRef\]](#)
2. Du, D.; Hwang, F.K.; Hwang, F. *Combinatorial Group Testing and Its Applications*; World Scientific: Singapore, 2000; Volume 12.
3. Farach, M.; Kannan, S.; Knill, E.; Muthukrishnan, S. Group testing problems with sequences in experimental molecular biology. In Proceedings of the Compression and Complexity of Sequences 1997, Salerno, Italy, 11–13 June 1997; pp. 357–367.
4. D'yachkov, A.G.; Polyanski, N.; Shchukin, V.Y.; Vorobyev, I. Separable Codes for the Symmetric Multiple-Access Channel. *IEEE Trans. Inf. Theory* **2019**, *65*, 3738–3750. [\[CrossRef\]](#)
5. Shental, N.; Levy, S.; Wuvshet, V.; Skorniakov, S.; Shalem, B.; Ottolenghi, A.; Greenshpan, Y.; Steinberg, R.; Edri, A.; Gillis, R.; et al. Efficient high-throughput SARS-CoV-2 testing to detect asymptomatic carriers. *Sci. Adv.* **2020**, *6*, eabc5961. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Gabrys, R.; Pattabiraman, S.; Rana, V.; Ribeiro, J.; Cheraghchi, M.; Guruswami, V.; Milenkovic, O. AC-DC: Amplification Curve Diagnostics for Covid-19 Group Testing. *arXiv* **2020**, arXiv:2011.05223.
7. Bui, T.V. A simple self-decoding model for neural coding. *bioRxiv* **2022**. [\[CrossRef\]](#)
8. D'yachkov, A.G.; Rykov, V.V. Bounds on the length of disjunctive codes. *Probl. Peredachi Informatsii* **1982**, *18*, 7–13.
9. Indyk, P.; Ngo, H.Q.; Rudra, A. Efficiently decodable non-adaptive group testing. In Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, Austin, TX, USA, 17–19 January 2010; pp. 1126–1142.
10. Porat, E.; Rothschild, A. Explicit nonadaptive combinatorial group testing schemes. *IEEE Trans. Inf. Theory* **2011**, *57*, 7982–7989. [\[CrossRef\]](#)
11. Ngo, H.Q.; Porat, E.; Rudra, A. Efficiently decodable error-correcting list disjunct matrices and applications. In Proceedings of the International Colloquium on Automata, Languages, and Programming, Zurich, Switzerland, 4–8 July 2011; pp. 557–568.
12. Cheraghchi, M. Noise-resilient group testing: Limitations and constructions. *Discret. Appl. Math.* **2013**, *161*, 81–95. [\[CrossRef\]](#)
13. Cheraghchi, M.; Nakos, V. Combinatorial group testing and sparse recovery schemes with near-optimal decoding time. In Proceedings of the 2020 IEEE Symposium on Foundations of Computer Science (FOCS), Virtual, 16–19 November 2020; pp. 1203–1213.
14. Cai, S.; Jahangoshahi, M.; Bakshi, M.; Jaggi, S. Efficient algorithms for noisy group testing. *IEEE Trans. Inf. Theory* **2017**, *63*, 2113–2136. [\[CrossRef\]](#)
15. Bondorf, S.; Chen, B.; Scarlett, J.; Yu, H.; Zhao, Y. Sublinear-time non-adaptive group testing with  $O(k \log n)$  tests via bit-mixing coding. *IEEE Trans. Inf. Theory* **2020**, *67*, 1559–1570. [\[CrossRef\]](#)
16. Price, E.; Scarlett, J. A Fast Binary Splitting Approach to Non-Adaptive Group Testing. In Proceedings of the Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM), Virtual, 17–19 August 2020.
17. De Bonis, A.; Vaccaro, U. Improved algorithms for group testing with inhibitors. *Inf. Process. Lett.* **1998**, *67*, 57–64. [\[CrossRef\]](#)
18. De Bonis, A.; Gasieniec, L.; Vaccaro, U. Optimal two-stage algorithms for group testing problems. *SIAM J. Comput.* **2005**, *34*, 1253–1270. [\[CrossRef\]](#)
19. Hwang, F.K.; Liu, Y. Error-tolerant pooling designs with inhibitors. *J. Comput. Biol.* **2003**, *10*, 231–236. [\[CrossRef\]](#)
20. Bui, T.V.; Kuribayashi, M.; Kojima, T.; Echizen, I. Sublinear decoding schemes for non-adaptive group testing with inhibitors. In Proceedings of the International Conference on Theory and Applications of Models of Computation, Kitakyushu, Japan, 13–16 April 2019; pp. 93–113.
21. Bui, T.V.; Chee, Y.M.; Scarlett, J.; Vu, V.K. Group Testing with Blocks of Positives. In Proceedings of the IEEE International Symposium on Information Theory, ISIT 2022, Espoo, Finland, 26 June–1 July 2022; pp. 1082–1087. [\[CrossRef\]](#)
22. Colbourn, C.J. Group testing for consecutive positives. *Ann. Comb.* **1999**, *3*, 37–41. [\[CrossRef\]](#)

23. Müller, M.; Jimbo, M. Consecutive positive detectable matrices and group testing for consecutive positives. *Discret. Math.* **2004**, *279*, 369–381. [[CrossRef](#)]
24. Juan, J.S.T.; Chang, G.J. Adaptive group testing for consecutive positives. *Discret. Math.* **2008**, *308*, 1124–1129. [[CrossRef](#)]
25. Bui, T.V.; Cheraghchi, M.; Nguyen, T.D. Improved algorithms for non-adaptive group testing with consecutive positives. In Proceedings of the 2021 IEEE International Symposium on Information Theory (ISIT), Melbourne, Australia, 12–20 July 2021; pp. 1961–1966.
26. Chang, H.; Tsai, Y.L. Threshold group testing with consecutive positives. *Discret. Appl. Math.* **2014**, *169*, 68–72. [[CrossRef](#)]
27. Chang, H.; Lan, W.C. Interval group testing for consecutive positives. *Discret. Math.* **2017**, *340*, 1488–1496. [[CrossRef](#)]
28. Kautz, W.; Singleton, R. Nonrandom binary superimposed codes. *IEEE Trans. Inf. Theory* **1964**, *10*, 363–377. [[CrossRef](#)]
29. Stinson, D.R.; Wei, R. Generalized cover-free families. *Discret. Math.* **2004**, *279*, 463–477. [[CrossRef](#)]
30. D'yachkov, A.; Vilenkin, P.; Torney, D.; Macula, A. Families of finite sets in which no intersection of  $\ell$  sets is covered by the union of  $s$  others. *J. Comb. Theory Ser. A* **2002**, *99*, 195–218. [[CrossRef](#)]
31. Chen, H.B.; Fu, H.L.; Hwang, F.K. An upper bound of the number of tests in pooling designs for the error-tolerant complex model. *Optim. Lett.* **2008**, *2*, 425–431. [[CrossRef](#)]
32. Ganesan, A.; Jaggi, S.; Saligrama, V. Non-adaptive group testing with inhibitors. In Proceedings of the ITW, Jerusalem, Israel, 26 April–1 May 2015; pp. 1–5.
33. Chang, H.; Chen, H.B.; Fu, H.L. Identification and classification problems on pooling designs for inhibitor models. *J. Comput. Biol.* **2010**, *17*, 927–941. [[CrossRef](#)] [[PubMed](#)]
34. Bruno, W.J.; Knill, E.; Balding, D.J.; Bruce, D.C.; Doggett, N.A.; Sawhill, W.W.; Stallings, R.L.; Whittaker, C.C.; Torney, D.C. Efficient pooling designs for library screening. *Genomics* **1995**, *26*, 21–30. [[CrossRef](#)]
35. Balding, D.J.; Torney, D.C. The design of pooling experiments for screening a clone map. *Fungal Genet. Biol.* **1997**, *21*, 302–307. [[CrossRef](#)]