Almost disjunct matrices from codes and designs

Arya Mazumdar*

Alexander Barg**

Abstract—A \((t, \epsilon)\) group testing scheme has the property of identifying any \(t\) defective subjects out of a population of size \(N\) while allowing \(\epsilon\) probability of false-positive for an item. We establish a new connection between \((t, \epsilon)\) schemes and error correcting codes based on the dual distance of codes, and construct schemes with \(M = O(t^{3/2} \log N)\) tests. We also outline a new connection between group testing schemes and combinatorial designs.

I. INTRODUCTION

Combinatorial group testing is an old and well-studied problem. In the most general form it is assumed that there is a set of \(N\) elements among which at most \(t\) are defective. This set of defective items is called a defective configuration. The smallest number of yes/no questions, or tests, that identify the defective configuration is \(\log \sum_{i=0}^t \binom{N}{i} \approx t \log \frac{N}{t}\). The main objective of the classical group testing problem is to identify the defective configuration using the number of tests that is as close to this minimum as possible.

In the group testing problem, a group of elements are tested together, and if this particular group contains any defective element, the test result is positive. Based on the test results of this kind one identifies (with an efficient algorithm) the defective configuration using the smallest possible number of tests. The collection of tests is called a group testing scheme. Group testing schemes can be adaptive (see, e.g., [3]), where the design of one test may depend on the results of preceding tests. In this paper we are interested in non-adaptive group testing schemes: here all the tests are designed together. If the number of designed tests is \(M\), then a non-adaptive group testing scheme is equivalent to the design of a binary test matrix of size \(M \times N\) where the \((i, j)\)th entry is 1 if the \(i\)th test includes the \(j\)th element; it is 0 otherwise. As the test results, we see the Boolean OR of the columns corresponding to the defective entries. It is known that the matrix must have disjunct property (defined later) to be a good group testing matrix. The best known lower bound on the number of required tests in terms of the number of elements \(N\) and the maximum number of defective elements \(t\) is given by [4],

\[ M = \Omega \left( \frac{t^2}{\log t} \log N \right). \]

The existence of non-adaptive group testing schemes with \(M = O(t^2 \log N)\) is also known for some time [3], [10], [18]. In contrast, for the adaptive setting, schemes have been constructed with the optimal number, \(O(t \log N)\), of tests [3], [9].

A construction of group testing schemes from error-correcting code matrices and using code concatenation [13] appeared in the seminal paper by Kautz and Singleton [11]. In [11], the authors concatenate a \(q\)-ary \((q > 2)\) Reed-Solomon code with a unit weight code to use the resulting codewords as the columns of the testing matrix. Generally, constructing groups testing schemes is a difficult problem. As one of the ways of attacking it, it has been suggested to construct schemes that permit a small probability of error (either missing defectives, or allowing false positives). Such schemes were considered under the name of weakly separated designs in [14], [15], [20] and independently in [12]. With this relaxation it is be possible to reduce the number of tests to be proportional to \(t \log N\) [20]; however, this result is not constructive. An explicit (nonprobabilistic) construction of weakly disjunct matrices with the number of tests proportional to \(O(t^{3/2} \log N)\) was presented in [16]. Related notions are also explored in [1], [5], [6] with explicit constructions.

The construction of [11] and many others are based on the so-called constant weight error-correcting codes, i.e., sets of binary vectors with a fixed number of ones. The link between the group-testing recovery property and codes that has been used in the existing works [11] relies on the minimum separation between the vectors of the code (the code distance). The main contribution of this paper is a refined analysis of constructions of group testing schemes that goes beyond the minimum distance and relates the number of tests \(M\) to the dual distance of the (constant weight) code. This allows us to connect weakly separated designs to error-correcting codes in a new and general way. Previously, apart from [16], the connection between error-correcting codes and weakly separated designs was known only for the very specific family of maximum distance separable codes [12], for which much more than the dual distance is known.

Based on the newfound connection, we provide some examples of explicit (constructible families) schemes of non-adaptive group testing that can identify all elements of a defective configurations of size at most \(t\), but allow a vanishing probability of false positive for each item. For instance, using algebraic-geometric codes, we construct explicit group testing schemes that groups of defective subjects of size \(t\) while allowing an \(\epsilon\) probability of false positives using \(O(t^{3/2} \log N)\) tests for \(\epsilon \to 0\).

We note a recent work [6] that suggests a way to construct weakly separated designs with \(O(t \text{poly}(\log N))\) tests by...
partitioning the subjects into blocks of equal parts and using optimal non-adaptive tests independently for each block. The approach of our work is completely different, with the main emphasis being placed on the connection between the disjunct property and dual distance of codes, which leaves a lot of room for studying specific constructions.

Finally, note that constant weight codes with a given value of dual distance \( d' \) are known as combinatorial designs (of strength \( d' - 1 \)). The use of \( t \)-designs for constructing disjunct matrices is not new, see, e.g., Sect. 7.4 of [3]. Special cases of designs have been used to construct nonadaptive group testing matrices for some particular parameters [19, Sec. 11.3]. However the conclusion in [3], p.146 is that disjunct matrices obtained from designs are of little interest because of restrictions on their parameters. Our approach is rather different from the one taken in the cited references, resulting in meaningful constructions of weakly disjunct matrices and group testing schemes.

II. DEFINITIONS AND NOTATIONS

Define the support of a vector \( \text{supp}(x), x \in (F_q)^n \) as the set of coordinates where \( z \) has nonzero entries. The support of a set of vectors \( X = \{x_i, i \geq 1\} \) is the union of supports \( \bigcup_{i \geq 1} \text{supp}(x_i) \).

Definition 1: An \( M \times N \) binary matrix \( A \) is called \( t \)-disjunct if the support of any of its columns is not contained in the union of the supports of any other \( t \) columns. It is easy to see that a \( t \)-disjunct matrix gives a group testing scheme that identifies any defective set up to size \( t \). Conversely, any group testing scheme that identifies any defective item outside of defective configuration as defective (false-positive).

Remark 1: Unless \( \epsilon < t/N \), the number of false positives in the \((t, \epsilon)\) scheme on average will be greater than the actual number of defectives. However even in that case, a vanishingly small proportioned subset will be provided as output that includes all of the \( t \) defective items.

Distance distribution of a code: We list a few concepts related to the distance distribution of codes. More information about them can be found, for instance, in [2].

1. Let \( C \subset (F_q)^M \) be a code. Define the distance distribution of \( C \) as the set of numbers \( A_0, A_1, \ldots, A_M \), where \( A_i = (1/|C|)\{\{(x, y) \in C^2 : d(x, y) = i\} \}, i = 0, 1, \ldots, M \) is the average number of ordered pairs of codevectors with Hamming distance \( i \). The smallest \( i \geq 1 \) such that \( A_0 > i \) is called the distance of the code \( C \), and we use the notation \( C(M, N, d) \) to refer to a code of length \( M \) and cardinality \( N \).

Define the dual distance \( d' \) of \( C \) as follows:

\[
d'(C) = \min \{ j > 1 : A'_j := \sum_{j=i}^{\infty} A_j K_j(i) > 0 \},
\]

where \( K_j(i) \) is the value of a Krawtchouk polynomial.

2. Similar concepts can be defined for constant weight codes. Let \( J_w^n \) be the set of all binary vectors with \( w \) ones and let \( C \subset J_w^n \) be a code. We use the notation \( C(M, N, d, w) \) to refer to a constant weight code of length \( M \), cardinality \( N \), and distance \( d \). Define the distance distribution of \( C \) as follows:

\[
b_i = \frac{1}{|C|} \{\{(x, y) \in C^2 : w - |\text{supp}(x) \cap \text{supp}(y)| = i\} \}
\]

for \( i = 0, 1, \ldots, w \). Define the dual distance \( d' \) of the constant weight code \( C \) as follows:

\[
d'(C) = \min \{ j > 1 : b'_j := \frac{1}{|C|} \sum_{i=0}^{w} b_i Q_j(i) > 0 \},
\]

where \( Q_j(i) \) is the value of a Hahn polynomial.

III. ALMOST DISJUNCT MATRICES FROM CODES

A. The Kautz-Singleton construction

The main observation behind this construction is the following result of Kautz and Singleton [11].

Proposition 2: An \( (M, N, d, w) \) constant weight binary code \( C \) provides a \( t \)-disjunct matrix, where \( t = \left\lfloor \frac{w-1}{w-d/2} \right\rfloor \).

Proof: Write the codewords of \( C \) as the columns of an \( M \times N \) matrix. The intersection of supports of any two columns has size at most \( w - d/2 \). Hence if \( w > t(w - d/2) \), support of any column will not be contained in the union of supports of any \( t \) other columns.

This proposition implies that a group testing scheme can be obtained from constant weight codes with large distance. [11] observed that such codes can be obtained from non-constant-weight \( q \)-ary codes in which every symbol is replaced by its binary indicator vector in the alphabet.

The main contribution of this paper is a refined analysis of the distance distribution of codes that gives rise to almost disjunct matrices (see Def. 2). The dual distance of the code defined above plays an important role in the analysis.
Our goal is to design an \( M \times N \) matrix \( A \) such that \( t \) randomly chosen columns do not contain the support of any other of its columns. As above, we form the matrix using the codewords of a constant weight code \( C(M, N, d, w) \) as the columns. Let the codewords of \( C \) be \( x_1, x_2, \ldots, x_N \) and let \( \mu_{ij} := d(x_i, x_j) \). Having in mind the design of almost disjoint matrices, we begin with the following extension of Prop. 2.

**Proposition 3:** The probability of violating the conditions of Def. 2 is bounded above as

\[
P_{R_t'}(\{I \in P_t(N), j \in [N] \setminus I : \text{supp}(c_j) \subseteq \cup_{k \in I} \text{supp}(c_k)\}) \leq t \Pr((k, j) \in [N]^2 : \mu_{jk} \leq 2(1 - \frac{1}{t})w) + O(1/N). \quad (4)
\]

**Proof:** Denote by \( P \) the probability on the left-hand side of (4). We argue as follows:

\[
P \leq P_{R_t'}(\{I \in P_t(N), j \in [N] \setminus I : \text{supp}(c_j) \subseteq \bigcup_{k \in I} \text{supp}(c_k)\})
\]

\[
\leq P_{R_t'}(\{I \in P_t(N), j \in [N] \setminus I : w \leq \sum_{k \in I} (w - \mu_{jk}/2)\})
\]

\[
\leq P_{R_t'}(\{I \in P_t(N), j \in [N] \setminus I : \exists k \in I, \mu_{jk} \leq 2(t - 1)w/t\})
\]

\[
\leq t \Pr(\{I \in P_t(N), j \in [N] \setminus I, k \in I : \mu_{jk} \leq 2(t - 1)w/t\})
\]

\[
= t \Pr(\{k \in [N], j \in [N] \setminus \{k\} : \mu_{jk} \leq 2(t - 1)w/t\}),
\]

where the probability on the last line is computed with respect to a pair of uniformly chosen distinct random vectors from \( C \). However, one can approximate this probability as one with respect to just a pair of uniformly chosen vectors from \( C \). Continuing the above calculation, we have,

\[
P \leq t \frac{N}{N-1} \Pr(\{(k, j) \in [N]^2 : \mu_{jk} \leq 2(1 - 1/t)w\})
\]

\[
\leq t \Pr(\{(k, j) \in [N]^2 : \mu_{jk} \leq 2(1 - \frac{1}{t})w\}) + O(1/N). \quad (5)
\]

Going forward we will neglect the \( O(1/N) \) term. We examine two different ways of constructing almost disjoint matrices from codes using this proposition.

**B. Case 1: Almost disjoint matrices from nonbinary codes**

In this section we estimate the probability in (4) for non-binary linear codes used in the Kautz-Singleton construction, beginning with an \( (n, N, d) \) \( q \)-ary code \( C \), where \( n = M/q \). To construct a group testing (almost disjoint) matrix we will map every symbol of the codeword to a binary vector of \( (q - 1) \) 0s and one 1 in the location that corresponds to the value of the symbol. We denote this mapping by \( \phi \). As a result, one obtains a set of binary vectors of length \( M \) and constant weight \( w = n = M/q \). The parameters of the resulting binary constant weight code are \( (M = qn, N, 2d, w = n = M/q) \).

The main result proved in this part is given by the following statement.

**Proposition 4:** Let \( M \) be an \( M \times N \) matrix constructed from a \( q \)-ary \( (n, N) \) code with dual distance \( d' \). For any even \( l < d' \) the probability in 4 is bounded above as follows:

\[
t \left(\frac{(q - 1)ne\ell}{2}\right)^{\ell/2} n^{-\ell} \left(\frac{q}{t} - 1\right)^{-\ell/2} \sum_{i=0}^{\ell/2} \left(\frac{(q - 1)\ell/2}{ne}\right)^{i}. \quad (6)
\]

Let \( Z \) be the random variable whose values are the distances between a pair of randomly and uniformly (with replacement) chosen codewords of the code \( C \). Note that

\[
\Pr(Z = i) = \frac{A_i}{N},
\]

where \( \{A_i, i = 1, \ldots, n\} \) is the distance distribution of the code \( C \). The following proposition is obvious.

**Proposition 5:** Let \( t < q \). The probability (4) is bounded above as follows:

\[
t \Pr(Z \leq (1 - 1/t)n) \leq \frac{\mathbb{E}[Z - (1 - \frac{1}{q})n]}{n^{\ell}(1 - \frac{1}{q})^{\ell}}, \quad (7)
\]

where \( \ell \) is an even number.

The moments on the right-hand side of (7) are estimated using the next two lemmas.

**Lemma 6:** Suppose we have a linear code \( C \) over \( \mathbb{F}_q \) of length \( n \) and size \( N \). Moreover assume \( d' \) is the dual distance of the code. For any \( \ell < d' \),

\[
\frac{1}{N} \sum_{j=0}^{n} \left( j - n \left(1 - \frac{1}{q}\right) \right)^\ell A_j
\]

\[
= \sum_{j=0}^{n} \left( j - n \left(1 - \frac{1}{q}\right) \right)^\ell \binom{n}{j} \left(1 - \frac{1}{q}\right)^{n-j}. \quad (8)
\]

**Proof:** See appendix.

Next we upper bound the right hand side of (8).

**Lemma 7:** Let \( 1/2 < p < 1 \). Then,

\[
\mu_n(2p) \leq \sum_{j=0}^{n} \left( \frac{j - np}{\sqrt{p(1-p)}} \right)^{2p} \binom{n}{j} p^{j} (1-p)^{n-j}
\]

\[
< (ner)^{\ell} \sum_{i=0}^{\ell} \frac{p^r}{(1-p)ne}^i.
\]

**Proof:** See appendix.

Using these two lemmas, and setting \( p = 1 - \frac{1}{q} \), we immediately have, for any even \( \ell < d^2 \) (the dual distance of the other code),

\[
\mathbb{E}[Z - (1 - \frac{1}{q})n]^\ell \leq \left(\frac{ne\ell}{2q}\right)^{\ell/2} \sum_{i=0}^{\ell/2} \left(\frac{(q - 1)\ell/2}{ne}\right)^{i}.
\]

Together with (4) this implies (6), finishing the proof.

Let us examine a few specific choices of outer \( q \)-ary codes. In each case our goal is to choose the parameters so that the quantity in (6) is small and examine the parameters of the resulting weakly group testing schemes.

1) **Reed-Solomon codes:** In this case, \( n = q - 1, \ell \approx \log_q N \). Let us take \( t = \frac{n}{q-1} \). Let us assume that Hence (6) is bounded above as

\[
\leq t \left(\frac{q}{2e}\right)^{\ell/2} (q/t - 1)^{-\ell/2} \sum_{i=0}^{\ell/2} \left(\frac{\ell}{2e}\right)^{i}.
\]
As long as \( r < d \), we obtain Solomon codes to codes on curves.

Better results in terms of the distance \( d \) when \( N > q \log_2 q \). As we have assumed, \( t = \frac{q}{t+1} \), and \( M = q(q-1) \), we have,

\[
M = O(t^2 \log_q N^2).
\]

1) Algebraic-geometric codes: Better results in terms of the number \( M \) of tests are obtained by moving from Reed-Solomon codes to codes on curves.

   a) Hermitian codes: Let \( q \) be a power of a prime and let \( 0 \leq r \leq q^3 \). There exists a family of linear \( q \)-ary codes of length \( n = q^3 \), cardinality \( N = q^2(r+1-(q-1)/2) \) with dual distance \( d' \geq r + q + 2 - q^2 \). In particular, choosing \( r = q^2 \), we obtain \( d' \geq q + 2 \). This suffices to ensure that the quantity \( (6) \) is small, i.e., the matrix formed by using Hermitian codes in the Kautz-Singleton construction is almost disjunct. Since we have assumed that \( t < q^2 \), we obtain \( M = q^2 \) for the number of tests. Finally, since \( N \approx (q^2)^r \), we observe that \( \log N = \Theta(q^2) \) and \( M = O(t^{3/2} \log_q N) \), matching the result of [16].

   b) Suzuki codes: Similar results are obtained if we take Suzuki codes, i.e., linear \( q \)-ary codes with \( q = 2q_n \) and \( q_0 = 2^{m} \) of length \( n = q^3 \), cardinality \( N = q^{r+1-q_0(q-1)} \) with dual distance \( d' \geq r - 2(q_0(q-1) - 1) \) [7]. Namely, taking \( r = 2q_0q \), we obtain \( d' \geq 2q_0 = (1/4)n^{1/4} \), which again ensures the almost disjoint property of the matrix obtained from the Kautz-Singleton construction. Observing that \( \log_q N \approx q_0 q = n^{3/4} \), we find that

\[
M = nq = O(n^{1/4} \log_q N) = O(t^{3/2} \log_q N).
\]

C. Case 2: Almost disjoint matrices from constant weight codes

In this part we study properties of matrices constructed from constant weight codes with a known value of the dual distance \( d' \). As is well known [2], a set of binary vectors with dual distance \( d' = r + 1 \) forms a combinatorial \( r \)-design (a collection of \( w \)-subsets of an \( n \)-set \( V \), called blocks, such that every \( t \) elements of \( V \) are contained in the same number of blocks).

Our plan is to use Prop. 3 in order to estimate the probability of a false positive. As a first step, we establish a result on moments of the distance distribution of constant weight codes.

Theorem 8: Let \( C \) be a constant weight code of weight \( w \), length \( n \), distance distribution \( \{b_i, i = 0, \ldots, w\} \) and dual distance \( d' \). Let \( X \) be a hypergeometric random variable with pmf and expectation

\[
f_X(i) = \binom{w}{i} \binom{n-w}{w-i}^{-1}, \quad i = 0, 1, \ldots, w; \quad \mathbb{E}X = \frac{w^2}{n}. \tag{9}
\]

As long as \( r < d' \),

\[
\sum_{i=0}^{w} \left( \binom{w}{n-w} - i \right) b_i = |C| \mathbb{E}[(X - \mathbb{E}X)^2]. \tag{10}
\]

Remark 2: It is known that the left-hand side of (10) is always greater than or equal to the right-hand side (this result is called the Sidel'nikov inequality. The remark about equality in (10) apparently is new.

For the lack of space we do not include the proof of (10) which relies on some simple properties of association schemes and applies both to constant weight codes and to the corresponding result for unrestricted code in (8).

Let \( x, y \) be two random vectors chosen from a constant weight code \( C(n, N, d, w) \) with replacement. Define the random variable \( Z = (1/2)d(x, y) \). The pmf of \( Z \) is given by \( p_Z(i) = b_i/N \), viz. (2). Now let us assume that \( t < n/w \), then the right-hand side of (5) equals

\[
t \mathbb{P} \left( Z \leq \frac{t-1}{t} w \right) = \mathbb{P} \left( \frac{w(n-w)}{n} - Z \geq \left( \frac{1}{t} - \frac{w}{n} \right) \right) \leq \mathbb{E} \left( \frac{w(n-w)}{n} - Z \right)^{\ell} \wedge \left( \frac{1}{t} - \frac{w}{n} \right)^{\ell} \wedge \left( \frac{1}{w} \right)^{\ell},
\]

for any even integer \( \ell \leq d' \), where \( X \) is given by (9).

It is evident that \( X = \sum_{i=1}^{w} X_i \), where \( X_1, \ldots, X_w \) denote values of random samples drawn without replacement from a population of \( n \) values with \( w \) ones and \( n-w \) zeros. Using a result of Hoeffding, [8], Thm. 4, we claim that

\[
\mathbb{E}[(X - \mathbb{E}X)^{\ell}] \leq \mathbb{E}[(Y - \mathbb{E}Y)^{\ell}],
\]

where \( Y \equiv \sum_{i=1}^{w} Y_i \), and \( Y_1, \ldots, Y_w \) denote samples from the same population with replacement. This makes the \( Y_i \)'s into independent Bernoulli(\( \frac{w}{n} \)) random variables, and

\[
\mathbb{E}[(X - \mathbb{E}X)^{\ell} = \sum_{j=0}^{w} (j - \frac{w^2}{n})^{\ell} \binom{w}{j} \left( \frac{w}{n} \right)^{j} \left( 1 - \frac{w}{n} \right)^{w-j} \\
= \sum_{j=0}^{w} (w - j) \binom{w}{j} \left( \frac{w}{n} \right)^{w-j} \left( 1 - \frac{w}{n} \right)^{j} \\
< 2 \left( \frac{we^\ell}{2} \right)^{\ell/2} \left( \frac{w}{n} \left( 1 - \frac{w}{n} \right) \right)^{\ell/2},
\]

where we have used Lemma 7 with \( p = 1 - w/n \) and assumed \( \ell(n-w) \leq 2w^2 \) (see the discussion after the proof of Lemma 7 in appendix).

Putting all this together, and assuming further \( 1/t > \frac{w}{n} + \sqrt{\ell} \) the right hand side of (5) is less than,

\[
2 \left( \frac{we^\ell}{2n(1/t - w/n)^2} \right)^{\ell/2} < 2^{\ell/2 + 1} = 0.
\]

From the assumptions on \( n, w, t \), it follows that, for the above to hold, \( n = 2.25t^2 \ell \) suffices. The factor \( \ell \) controls how fast the probability \( P \) goes to zero, but large \( \ell \) increases the number of samples \( M = n = O(t^2 \ell) \). It remains to use results about the existence of \( t \)-designs. One of the first such results claims that for any \( \ell \) there exists an \( \ell \)-design with \( N = \left( \frac{n}{\ell} \right)^{\ell} \) [19, Thm. 9.8]. Our current research involves finding smaller-sized designs for the cases of our interest.
The lemma follows immediately.

Proof of Lemma 7: Let us define i.i.d. random variables $X_i$, $i = 1, \ldots, n$ in the following way:

$$X_i = \begin{cases} \sqrt{(1-p)/p} & \text{with probability } p \\ -\sqrt{p/(1-p)} & \text{with probability } 1-p. \end{cases}$$

Note that $E X_i = 0$ and $E X_i^2 = 1$. Let $X = \sum_{i=1}^n X_i$. Clearly, with probability $\left(\binom{n}{i} p^i (1-p)^{n-i}\right)$

$$X = i\sqrt{(1-p)/p} - (n-i)\sqrt{p/(1-p)} = \frac{i - np}{\sqrt{p(1-p)}}.$$  

Hence,

$$\mu_n(2r) = E X^{2r}.$$  

To evaluate $E X^{2r}$ we follow a method outlined by Terence Tao. We have

$$E X^{2r} = \sum_{i_1, \ldots, i_{2r}} E X_{i_1} \cdots X_{i_{2r}}.$$  

We use the fact that the variables $X_i$ are independent. If at least one of the $X_i$ appears only once then the corresponding monomial is zero. So it may be assumed that each index appears at least twice in the expectations that contribute to the sum. In particular, there are at most $r$ distinct $X_i$’s that can appear. Suppose that $r - t$ such terms appear. Then, using the fact that $X_i$’s have unit variance and $|X_i| \leq \sqrt{p/(1-p)}$, we have

$$E X^{2r} \leq \sum_{t=0}^r N_t \left(\frac{\sqrt{p}}{\sqrt{1-p}}\right)^{2t} = \sum_{t=0}^r N_t \left(\frac{p}{1-p}\right)^t,$$

where $N_t$ is the number of ways one can assign integers $i_1, \ldots, i_{2r} \in \{1, \ldots, n\}$ such that each $i_j$ appears, at least twice and exactly $r - t$ integers appear.

A crude bound on $N_t$ gives,

$$N_t \leq \left(\frac{n}{r-t}\right)^{(r-t)2^t} \leq \left(\frac{ne}{p-t}\right)^{(r-t)2^t}.$$  

Hence,

$$E X^{2r} \leq \sum_{t=0}^r \left(\frac{ne}{p-t}\right)^{(r-t)2^t} \left(\frac{p}{1-p}\right)^t\frac{1}{(1-p)^t} \leq (ner)^r \sum_{t=0}^r \left(\frac{pr}{ne(1-p)}\right)^t.$$  

When $\frac{p}{1-p} \leq \frac{n}{r}, E X^{2r}$ in the above proof can be further simplified. Namely, $\frac{pr}{ne(1-p)} \leq e^{-1}$, and we have,

$$E X^{2r} \leq \frac{1}{1-e^{-1}}(ner)^r.$$  

References


