On the impossibility of amplifying the independence of random variables

(Extended Abstract)

Jin–yi Cai* Suresh Chari†

Abstract

In this paper we prove improved lower and upper bounds on the size of sample spaces which are required to be independent on specified neighborhoods. Our new constructions yield sample spaces whose size is smaller than previous constructions due to Schulman[8]. Our lower bounds generalize the known lower bounds of [1, 3, 2]. In obtaining these bounds we examine the possibilities and limitations of amplifying limited independence by fixed functions. We show that in general independence can not be amplified from k–wise independence to $(k+1)$–wise independence. Finally, we enumerate all possible logical consequences of pair–wise independent random bits.

1 Introduction

Derandomization has proved to be an essential tool over the last few years in obtaining deterministic algorithms both in the sequential and parallel domain. For several problems the only deterministic algorithms have been obtained by derandomization. A very useful technique in derandomization is the use of *limited independence* of random variables.

Definition 1 *Random variables $X_1, X_2, \ldots, X_n$ are k–wise independent if for any set $I \subseteq \{1, \ldots, n\}$ of at most k indices and for any choice $v_1, \ldots, v_n$ we have*

$$\Pr[\wedge_{i\in I}(X_i = v_i)] = \prod_{i\in I} \Pr[X_i = v_i]$$

In this paper we will be primarily concerned with random bits, *i.e.,* uniform binary random variables. Efficient constructions are given in [7, 5, 1] for sample spaces $S \subseteq \{0,1\}^n$ such that the distribution induced on $\{0,1\}^n$, by sampling uniformly from S, is k–wise independent. These constructions have been used to design several deterministic algorithms from their randomized counterparts such as the NC algorithms for maximal independent set and other problems[7, 1].

*Dept. of Computer Science, State Univ. of New York at Buffalo, Buffalo NY 14260. Email: cai@cs.buffalo.edu. Supported in part by NSF grant CCR 9057486 and CCR 9319093, and an Alfred P. Sloan Fellowship.

†Dept. of Computer Science, Cornell Univ., Ithaca NY 14853. Email: chari@cs.cornell.edu. Supported in part by NSF grant CCR-9123730.

It can be shown that sample spaces defining a k–wise independent distribution must contain at least $\binom{n}{1} + \ldots + \binom{n}{\lfloor k/2 \rfloor}$ points[1, 2, 3]. Thus, polynomial size constructions are possible only for constant k. Schulman[8] observes that for several algorithms we do not need independence in all the $\binom{n}{k}$ subsets of indices and constructs sample spaces which are independent only for specified index sets of k elements. He gives constructions of sample spaces which are independent only on specified index sets and uses these constructions to get efficient algorithms for several problems.

In this paper we study the question of how large sample spaces must be in order to be independent on specified sets of index sets. For pair–wise independence the input constraints can be specified as a graph where an edge between i and j implies that we want independence between X_i and X_j. The best bounds known prior to this work for a sample space S satisfying the constraints specified by G were: $\alpha(G) \leq |S| \leq \Delta(G)$[8, 1, 2, 3], where $\alpha(G)$ is the clique number and $\Delta(G)$ is the maximum degree of G. We improve the lower bound to show that $|S| \geq \nu(G)$, where $\nu(G)$ is the function defined by Lovász [6] to approximate the Shannon capacity of graphs. We also show that there exists a sample space of size $\chi(G)$ where $\chi(G)$ is the chromatic number of G. Since we can approximate $\chi(G)$ by the maximum degree $\Delta(G)$ and there is an easy algorithm to color G with $\Delta(G)+1$ colors, we can improve the construction in [8] by constructing in NC a sample space of size smaller than $|E(G)|$, as given in [8].

For the general case of k–wise independence for $k > 2$ we use a standard reduction to the pair–wise case to obtain a lower bound which generalizes the previously known lower bounds. To obtain good upper bounds using this reduction we come against the following problem: Given n binary random variables $X_1, \ldots, X_n$ can we define 3 functions f_1, f_2, and f_3 on these variables such that, for any distribution $\mathcal{D}$ which guarantees that the X_i's are pair–wise independent, the three new random variables $f_j(X_1, \ldots, X_n)$ are three-wise independent? Note that if n is small then we can easily answer the question in the negative. When n is large, however, there will be enough entropy in the system, so that such a construction is conceivable. In this paper we show using a novel proof technique that irrespective of how large n is, we can construct a distribution $\mathcal{D}_0$ such that no fixed functions can yield three–wise independence. The proof simultaneously uses a p-adic reduction as well as volume and dimension estimates in a certain high dimensional space. Our proof technique also directly extends to show that no fixed functions can amplify k–wise independence into $(k+1)$–wise independence.

We then consider the question of what the logical consequences of pair–wise independence are, *i.e.:* what kind of events can fixed functions guarantee to occur over all distributions. For the pair–wise case we enumerate all the logical consequences. This paper is organized as follows: In section 2 we describe our results on new bounds for sample spaces independent on specified neighborhoods. In section 3 we prove a general result on the non–amplifiability of independence. Finally, in section 4 we enumerate the logical consequences of pair–wise independence.

2 Bounds for sample spaces uniform on neighborhoods

In this section we describe our results on new bounds for sample spaces which are independent on specified index sets. This approach was first considered by Schulman[8] who observed that for several algorithms we do not need independence in all the $\binom{n}{k}$ subsets of indices and constructs sample spaces which are independent only for specified index sets of k elements. A *neighborhood* is a set of indices for which the independence constraint is to be satisfied. The main result in [8] is

Theorem 2 *Given neighborhoods $N_1, \ldots, N_m \subseteq \{1, \ldots, n\}$ each of size at most k with each $i \in \{1, \ldots, n\}$ occurring in at most d neighborhoods, there exists an explicitly constructible sample space of size $O(d2^k)$ with the induced distribution independent on each of the specified neighborhoods.*

We are interested in the problem of establishing tight bounds on the size of the sample spaces independent on given neighborhoods.

Definition 3 *$S(n, k, m, d)$ denotes the size of the smallest sample space on n variables independent on m specified neighborhoods of size at most k, with each index occurring in at most d neighborhoods.*

In our notation, Theorem 2 says that $S(n, m, k, d) = O(d2^k)$. Consider the case when $k = 2$. Here, the input can be specified as a graph on n vertices where an edge $\{i, j\}$ implies that independence is to be satisfied over the set $\{i, j\}$. For $k = 2$, Theorem 2 gives a sample space of size $O(\Delta)$ where Δ is the maximum degree of the input graph. This is not optimal as shown by the following.

Lemma 4 *$S(n, 2, m, d) = O(\chi(G))$, where $\chi(G)$ is the chromatic number of the input graph G.*

Proof. Fix color classes $C_1, \ldots, C_{\chi(G)}$ in G. Then construct sample space S' which is pairwise independent over all pairs on $\chi(G)$ new variables $Y_1, \ldots, Y_{\chi(G)}$. Sample space S is defined by setting X_j equal to Y_k where the color class of the vertex j is C_k. Whenever vertices i and j are adjacent they belong to different color classes and hence the corresponding variables X_i and X_j are pairwise independent. We can explicitly construct such an S' with $|S'| = |S| = O(\chi(G))$[1, 5, 7]. $\square$

If G is a bipartite graph the above sample space is of constant size where as the construction in [8] has size $O(\Delta)$. Thus, for $k = 2$, Schulman's constructions essentially approximate the chromatic number by the maximum degree. Since G can be colored in NC with $\Delta + 1$ colors[7] we get

Corollary 5 *There exists an explicitly constructible (in NC) sample space of size $O(\Delta)$.*

This is smaller than the NC constructions of[8] which are of size $O(|E|)$.

$\chi(G)$ is not a lower bound as shown by the following class of graphs: Let $G(n) = (V, E)$, where $V = \{w \in \{-1, 1\}^n |\ w \text{ has } \frac{n}{2} \text{ ones } \}$ and with an edge between u and v if the dot product $u^T v$, over $\Re$, is 0. The following is a corollary of a result of [4].

Theorem 6 *If $n = 4p^k$, where p is an odd prime, then all independent sets in $G(n)$ are of size at most $2\binom{n-1}{\frac{n}{4}-1}$ and hence $\chi(G(n))$ is exponential in n.*

However there is a sample space, S, of size n, which satisfies all the independence constraints specified by $G(n)$: for $v \in V$, let $X_v = \frac{(1-v_i)}{2}$ at the i^{th} point in S, where v_i is the i^{th} component of v. By definition, vertices u and v are adjacent in $G(n)$ iff there are exactly $\frac{n}{4}$ positions i, such that $u_i = v_i = 1$. Thus, when u and v are adjacent, if we uniformly sample from S, we have $\Pr[(X_u = 0) \wedge (X_v = 0)] = \frac{1}{4}$. Thus, S satisfies all the pairwise independence constraints specified by $G(n)$.

We now show that a lower bound for $S(n, 2, m, d)$ is provided by the following interesting function of Lovász [6]. In 1979, Lovász introduced the function $\nu(G)$, in his study of the Shannon capacity of a graph:

Definition 7 *An orthonormal representation of a graph G on n vertices is a system of unit vectors $(u_1, \ldots, u_n)$, $u_i \in \Re^d$, such that if i and j are adjacent then u_i and u_j are orthogonal. The* value *of the representation is* $\min_c \max_{1 \leq i \leq n} \frac{1}{(c^T u_i)^2}$ *where c ranges over all unit vectors. $\nu(G)$ is defined to be the minimum value over all representations of G.*

Theorem 8 *$S(n, 2, m, d) \geq \nu(G)$, where G is the input graph of independence constraints.*

Proof. (Sketch) A sample space S which satisfies pairwise independence constraints given by a graph G yields an orthonormal representation for G: Let u_i be obtained by normalizing the vector whose j^{th} component is -1 if X_i is 1 at the j^{th} point of S and 1 otherwise. For any two vertices i and j, the dot product $u_i^T u_j$ is proportional to the number of components at which they have the same value minus the number of components at which they disagree. If i and j are adjacent, the corresponding variables X_i and X_j are pairwise independent and hence $u_i^T u_j = 0$. Thus, $(u_1, \ldots, u_n)$ is an orthonormal representation for G with each vector of dimension $|S|$. The lower bound follows from the fact that if G admits a dimension d orthonormal representation then $\nu(G) \leq d$ [6]. □

Now consider the case $k > 2$. Let $N_1, \ldots, N_m$ be the neighborhoods over which we want independence. Construct a graph $G_{n,k} = (V, E)$ with $V = \{I \subseteq \{X_1, \ldots, X_n\} \mid |I| \leq k\}$ and $E = \{(I, I') \mid I \cup I' \subseteq N_j \text{ for some } j\}$. Now, any sample space S which satisfies independence on the neighborhoods N_j, yields a sample space of the same size which satisfies the pairwise constraints of $G_{n,k}$: for $I \subseteq \{X_1, \ldots, X_n\}$ define $Y_I = \oplus_{i \in I} X_i$. Since S satisfies all the specified independence constraints, by the definition of k-wise independence, the Y_I's satisfy the pairwise constraints of $G_{n,k}$. Thus a lower bound for $S(n, k, m, d)$ is $\nu(G_{n,k})$. This generalizes the lower bounds of [1, 2, 3] since if we need independence in all k sized neighborhoods, then $G_{n,k}$ has a clique of size $\binom{n}{1} + \ldots + \binom{n}{\lfloor k/2 \rfloor}$ and for all graphs G, $\nu(G) \geq \omega(G)$ [6], where $\omega(G)$ is the clique number of G. The next main question we address in this paper is whether the converse is true, *i.e.:*

> Can we define n functions on the variables associated with $G_{n,k}$ such that given any sample space S satisfying the constraints of $G_{n,k}$ the induced distribution on the n new variables satisfies the given $(k+1)$–wise independence constraints?

3 Independence can not be amplified

In this section we consider *logical consequences* of pairwise independence of n binary random bits and, specifically, whether any fixed functions of pairwise independent bits could yield higher independence.

Definition 9 *An event E occurring with probability p is a* logical consequence *of pairwise independence of random bits $X_1, \ldots, X_n$ if for every distribution $\mathcal{D}$ which is pairwise independent on the X_i's the event E occurs with probability p.*

We show that there is no event E such that E occurs with probability $\frac{1}{8}$ is a logical consequence of pairwise independence of n bits, for all n. This implies that there are no fixed functions on n pairwise–independent bits which are 3-wise independent, thus answering the question in the end of the previous section in the negative.

Theorem 10 *Let $X_1, \ldots, X_n$ be uniform binary random variables. There exists a sample space $S \subseteq \{0,1\}^n$, such that, the induced distribution on the X_i's is unbiased and* pairwise independent, *but for any fixed functions $Y_j = f_j(X_1, \ldots, X_n)$, $1 \leq j \leq 3$, the random variables Y_1, Y_2 and Y_3 are not 3–wise independent unbiased random bits.*

Proof. In the proof, elements of $\{0,1\}^n$ are viewed as subsets of $\{1, \ldots, n\}$ where an n–bit vector is treated as the characteristic vector of the subset it represents. A pairwise–independent distribution must satisfy three sets of constraints: $\sum_\alpha p_\alpha = 1$, $\sum_{\alpha : i \in \alpha} p_\alpha = \frac{1}{2}$ for $1 \leq i \leq n$, and $\sum_{\alpha : i,j \in \alpha} p_\alpha = \frac{1}{4}$ for $1 \leq i \neq j \leq n$, where p_α is the probability assigned to point $\alpha \in \{0,1\}^n$. Conversely, any assignment p_α, $\alpha \in \{0,1\}^n$, which satisfies these constraints along with the condition that $p_\alpha \geq 0$ for all α, defines a distribution which is pairwise independent. We show that if event E occurring with probability p is a logical consequence of pairwise independence then p must be an integer combination of $1, \frac{1}{2}$, and $\frac{1}{4}$.

We rewrite the above constraints as the linear system of equations $\sum_{\alpha \in \{0,1\}^n} a_{S,\alpha} p_\alpha = b_S$ for $S = \emptyset, \{1\}, \ldots, \{n\}, \{1,2\}, \{1,3\}, \ldots, \{n-1, n\}$ where

$$a_{S,\alpha} = \begin{cases} 1 & \text{if } S \subseteq \alpha \\ 0 & \text{otherwise} \end{cases}$$

Let A be the matrix of coefficients of this linear system. We index the rows of A by the subset S that defines the row. Thus a pairwise independent distribution must satisfy the linear system $Ap = b$. First, we observe that the matrix A is of full row rank *over all fields.* To see this, consider columns of A defined by the elements of $\{0,1\}^n$ corresponding to the following subsets of $\{1, \ldots, n\}$: $\emptyset, \{1\}, \ldots, \{n\}, \{1,2\}, \ldots, \{1,n\}$ $\{2,3\}, \ldots, \{n-1,n\}$. The resulting sub–matrix is a square matrix which is upper triangular with 1's on the diagonal and hence invertible over all fields. Thus, A is of rank $1 + \binom{n}{1} + \binom{n}{2}$ over all fields.

Assume that there exist three functions f_1, f_2 and f_3, such that for any distribution which satisfies the pairwise–independence constraints for the X_i's , the random variables $Y_j = f_j(X_1, \ldots, X_n)$, $1 \leq j \leq 3$ are independent and unbiased random bits. Consider the event $E = [(Y_1 = 1) \wedge (Y_2 = 1) \wedge (Y_3 = 1)]$ which occurs with probability $\frac{1}{8}$ if the Y_j's are independent. Construct the 2^n dimensional $\{0,1\}$ vector $\vec{v}$ with $\vec{v}_\alpha = 1$ if E occurs at α. By assumption, any nonnegative solution p of the system $Ap = b$, also satisfies $\vec{v}p = \frac{1}{8}$ since p defines a pairwise independent distribution. We claim that this can happen only if $\vec{v}$ is in the span of the rows of A.

To see this, first fix p^* to be the vector with each entry $\frac{1}{2^n}$, which is a nonnegative solution of the system $Ap = b$ as it corresponds to the unbiased n–wise independent distribution. Let $z_1, \ldots, z_d$ be a basis for the null space of A, where $d = 2^n - 1 - \binom{n}{1} - \binom{n}{2}$. All solutions to the linear system $Az = b$ are of the form $p^* + \sum_{i=1}^{d} x_i z_i$, where x_i ranges over the underlying field Q. For small x_i the point $p^* + x_i z_i$ is in $(0,1)^{2^n}$ as the unit cube $(0,1)^{2^n}$ is an open set. Therefore $p^* + x_i z_i$ defines a probability distribution which satisfies $Ap = b$ and hence is a pair–wise independent distribution and so, by assumption, $v(p^* + x_i z_i) = \frac{1}{8}$. Since $vp^* = \frac{1}{8}$ as well, $v \cdot z_i = 0$ for all i and thus, v is in the linear span of the rows of A. Therefore we can write $\vec{v}$ as a linear combination of the rows, $A_\emptyset, \ldots, A_{\{n-1,n\}}$, of A with rational coefficients. Multiplying by the least common multiple of the denominators, we have

$$c\vec{v} = \sum_S c_S A_S$$

where the c_S's and c are integers with $gcd(c, c_\emptyset, \ldots c_{\{n-1,n\}}) = 1$.

We claim that $c = 1$. If not then consider a prime q which occurs in c and reduce the above equation modulo q. It follows that the rows of A are *linearly dependent* over the finite field Z_q which contradicts the fact that the matrix A has full row rank over all fields. Therefore $c = 1$ and $\vec{v}$ is an integer combination of the rows of A.

By assumption, if $Ap = b$ with p nonnegative then we have $\vec{v}p = \frac{1}{8}$ and since $\vec{v}$ is an integer combination of the rows of A, this implies that $\frac{1}{8}$ is an integer combination of $1, \frac{1}{2}$, and $\frac{1}{4}$ which is a contradiction.

We showed that the (affine) hyperplane $v \cdot p = \frac{1}{8}$ does not contain the portion of the (affine) d–dimensional solution space $Ap = b$ inside the unit cube $(0,1)^{2^n}$. Thus the intersection of the hyperplane with the d-dimensional solution space $Ap = b$ is a $(d-1)$–dimensional affine space. Excluding such points on the lower dimensional space of the intersection, which has d–dimensional volume 0, we can find points in the unit cube $(0,1)^{2^n}$ which satisfy the pairwise independence constraints and yet the variables Y_1, Y_2, Y_3 are not independent. Note that the normal vector v to the hyperplane $vp = \frac{1}{8}$ was defined by the event $[(Y_1 = 1) \wedge (Y_2 = 1) \wedge (Y_3 = 1)]$ and there are only finitely many such v corresponding to different choices for f_1, f_2, and f_3. Thus, in fact, we can find a pair–wise independent distribution p which is on none of the finitely many hyperplanes $v \cdot p = \frac{1}{8}$, and thus for no choice of the functions f_1, f_2 and f_3 can the variables Y_1, Y_2, Y_3 be independent and unbiased random bits. (In fact, in a measure theoretic sense, "almost all" pair–wise independent distributions p satisfy this.) □

What if we do not require the variables Y_1, Y_2, Y_3 to be necessarily unbiased? In fact the above proof shows that as long as none of Y_1, Y_2, Y_3 is a trivial random variable, i.e., a constant 0 or 1 valued random variable, the same conclusion still follows. (For trivial random variables, the conclusion is certainly false, as they are always "independent").

To see this, notice that $\sum_{i_1,i_2,i_3=0,1} \Pr[(Y_1 = i_1) \wedge (Y_2 = i_2) \wedge (Y_3 = i_3)] = 1$. Thus, for some i_1, i_2, i_3, $\Pr[(Y_1 = i_1) \wedge (Y_2 = i_2) \wedge (Y_3 = i_3)] \leq \frac{1}{8}$. But it can't equal 0, otherwise, by independence, we would have $\Pr[(Y_j = i_j] = 0$ for some j, contrary to our hypothesis. Now the same proof goes through, since no positive quantity $\leq \frac{1}{8}$ can be an integer combination of $1, \frac{1}{2}$ and $\frac{1}{4}$.

Corollary 11 *Let $X_1, \ldots, X_n$ be uniform binary random variables. Then there is a sample space $S \subseteq \{0,1\}^n$, such that, the induced distribution on the X_i's is unbiased and* pairwise independent, *but for any functions $Y_j = f_j(X_1, \ldots, X_n)$, $1 \leq j \leq 3$, the variables Y_1, Y_2 and Y_3 are not 3–wise independent, unless one of the Y_j's is constant.*

The proof of the theorem and the corollary directly generalize to the case $k > 2$.

Theorem 12 *For any k and n, let $X_1, \ldots, X_n$ be uniform binary random variables. There exists a sample space $S \subseteq \{0,1\}^n$, such that, the induced distribution on the X_i's is unbiased and k–wise independent, but for any fixed functions $Y_j = f_j(X_1, \ldots, X_n)$, $1 \le j \le (k+1)$, the random variables $Y_1, \ldots, Y_{k+1}$ are not $(k+1)$–wise independent unless one of the Y_j's is a constant.*

4 Enumerating all logical consequences

We turn to a slightly general problem: Let $X_1, \ldots, X_n$ be binary random variables with $\Pr[X_i = 1] = p_i$. We are interested in enumerating all possible logical consequences of pairwise independence of these variables. From the proof of Theorem 10, the characteristic vector of any event E that is a logical consequence must be an integer combination of the rows of the constraint matrix A. Furthermore, not all integer combinations are allowed: the resulting vector, which is the characteristic vector of E, must have $0, 1$ entries. This additional constraint allows us to enumerate all possibilities.

Theorem 13 *The only logical consequences of pairwise independence of n binary random variables $X_1, \ldots, X_n$ with $\Pr[X_i = 1] = p_i$, are events with probabilities 0, p_i, p_ip_j, $p_i(1-p_j)$, $(1-p_i)(1-p_j)$, $p_i(1-p_j)+p_jp_k$, $p_i(1-p_j)+p_j(1-p_i)$, $p_i+p_j+p_k-p_ip_j-p_ip_k-p_jp_k$, $p_i+p_j-p_ip_j-p_jp_k$, $p_i+p_j-p_ip_j-p_ip_k-p_lp_j+p_lp_k$ and $p_ip_j+p_k(1-p_j)+(1-p_k)(1-p_i)$ and their complements, where i, j, k and l are distinct indices.*

Proof. As in the previous section we index the rows of the constraint matrix A by the subset which defines the row of the matrix. In any integer combination of the rows of A, we denote the coefficient of the row $A_\emptyset$ as c_0, the coefficient of the row $A_{\{i\}}$ as c_i, $1 \le i \le n$ and the coefficient corresponding to $A_{\{i,j\}}$ as c_{ij}, $1 \le i \ne j \le n$. Also, elements of $\{0,1\}^n$ are viewed as subsets of $\{1, \ldots, n\}$ and we fix some arbitrary ordering of these subsets and use it to index 2^n dimensional vectors. Thus, for a 2^n dimensional vector $\vec{u}$, $\vec{u}_S$ denotes the entry of $\vec{u}$ corresponding to subset S.

Consider an event E whose characteristic vector, $\vec{v}$, is an integer combination of the rows of A. The condition that the integer combination result in a 0,1–vector forces restrictions on the coefficients which we outline below. We repeatedly use the fact that for any $S \subseteq \{1, \ldots, n\}$, $\vec{v}_S = c_0 + \sum_{i \in S} c_i + \sum_{i<j \in S} c_{ij}$. Assume, without loss of generality, that $\vec{v}_\emptyset = 0$, for if not, we can consider the complement of the event E. Thus, $c_0 = 0$ and $\vec{v}_{\{i\}} = c_i$ for all i and hence, c_i must be 0 or 1.

Claim 14 *At most three of the coefficients c_i, $1 \le i \le n$ can be 1.*

Proof. Since $\vec{v}_{\{i_1,i_2\}} = c_{i_1} + c_{i_2} + c_{i_1 i_2}$, if both c_{i_1} and c_{i_2} are 1 then $c_{i_1 i_2}$ must be -1 or -2. If four different coefficients c_{i_j}, $1 \le j \le 4$, are 1 then $\vec{v}_{\{i_1,i_2,i_3,i_4\}} = \sum_j c_{i_j} + \sum_{j<k} c_{i_j i_k} \le -2$. □

If c_i and c_j are both 0 then since $\vec{v}_{\{i,j\}} = c_i + c_j + c_{ij}$, the coefficient c_{ij} must be 0 or 1.

Claim 15 *There is at most one pair $\{i, j\}, i \ne j$ such that $c_i = c_j = 0$ and $c_{ij} = 1$.*

PROOF. Assume that there are 2 such pairs $\{i_1, i_2\}$ and $\{i_3, i_4\}$. Since some of the indices may be equal the set $S = \{i_1, i_2\} \cup \{i_3, i_4\}$ may contain either 3 or 4 elements. By definition $\vec{v}_S = \sum_{i_k \in S} c_{i_k} + \sum_{i_k < i_l \in S} c_{i_k i_l}$. For any pair $\{i_k, i_l\}$ in S, $i_k \neq i_l$, since $c_{i_k} = c_{i_l} = 0$ the coefficient $c_{i_k i_l}$ is either 0 or 1 and there are at least 2 different pairs with this coefficient 1. Thus, $\vec{v}_S \geq 2$ which is a contradiction and hence there can be at most one such pair. □

If $c_i = 1$ and $c_j = 0$, the coefficient c_{ij} should be either 0 or -1. In the next two claims we establish properties of such pairs.

Claim 16 *There are at most 2 pairs $\{i, j\}, i \neq j$ such that $c_i = 1$, $c_j = 0$ and $c_{ij} = -1$. If there are 2 such pairs $\{i_1, j_1\}$ and $\{i_2, j_2\}$ then $j_1 \neq j_2$ and $c_{j_1 j_2} = 1$.*

PROOF. Assume that there are three pairs $\{i_1, j_1\}$,$\{i_2, j_2\}$ and $\{i_3, j_3\}$ satisfying the above conditions. To handle the case of some of the coefficients being the same we first consider the set $I = \{i_1\} \cup \{i_2\} \cup \{i_3\}$. If $i_k, i_l \in I$ with $i_k \neq i_l$ then, since c_{i_k}=c_{i_l}=1, the coefficient $c_{i_k i_l}$ is either -1 or -2. Similarly consider the set $J = \{j_1\} \cup \{j_2\} \cup \{j_3\}$. For any two indices $j_k, j_l \in J$, $j_k \neq j_l$, the coefficient $c_{i_k i_l}$ is either 0 or 1. However by Claim 15 at most one such coefficient can be 1. If $i_k \in I$ and $j_l \in J$ then the coefficient $c_{i_k j_l}$ is either 0 or -1 and by assumption there are at least three such pairs with coefficient -1. Finally, let $S = I \cup J$. From the previous statements we can deduce that $\vec{v}_S \leq |I| - \binom{|I|}{2} + 1 - 3 \leq -1$ which is a contradiction.

Assume that 2 pairs $\{i_1, j_1\}$ and $\{i_2, j_2\}$ exist that satisfy the conditions of the claim. Let $S = \{i_1, j_1\} \cup \{i_2, j_2\}$. If $j_1 = j_2$ then $i_1 \neq i_2$ and since $c_{i_1 i_2} \leq -1$, we have $\vec{v}_S \leq -1$. Thus $j_1 \neq j_2$. Again if $c_{j_1 j_2} = 0$ we can show that $\vec{v}_S < 0$ irrespective of whether $i_1 = i_2$ or not. □

Claim 17 *If $c_i = 1$ and $c_{jk} = 1$ then either $c_{ij} = -1$ or $c_{ik} = -1$.*

PROOF. Firstly, since c_j and c_k are both nonnegative if $c_{jk} = 1$ then both c_j and c_k must be 0. In particular, i, j, k are all distinct. Hence, if both the coefficients c_{ij} and c_{ik} are 0, then $\vec{v}_{\{i,j,k\}} = 2$. □

From Claim 14 at most three of the c_i's are 1. Let N denote the number of c_i's which are one. We consider each case separately.

N=0: By Claim 15 the event E must have probability either 0 or $p_i p_j$ for some $i < j \leq n$.

N=1: Let $c_i = 1$. Then by Claim 16 at most 2 of the coefficients c_{ij} can be -1. If there are 2 such coefficients $c_{ij} = c_{ik} = -1$ then $c_{jk} = 1$. The other claims now imply that no other coefficient can be non zero and the event E has probability $p_i - p_i p_j - p_i p_k + p_j p_k$ which is $1 - [p_i p_j + (1 - p_j) p_k + (1 - p_i)(1 - p_k)]$. If exactly one $c_{ij} = -1$ then by Claims 15 and 17 the resulting events must have probabilities $p_i - p_i p_j$, $p_i - p_i p_j + p_j p_k$. If no $c_{ij} = -1$ then by Claim 17 the event E has probability p_i.

N=2: Let $c_i = c_j = 1$ with $i \neq j$. Then the coefficient c_{ij} must be either -1 or -2. If $c_{ij} = -2$, we argue that all other coefficients must be 0. If either $c_{ik} = -1$ or c_{jk} is -1 for some $k \neq i, j$, then, by Claim 16 exactly one of them is -1. And we get a contradiction since then $\vec{v}_{\{i,j,k\}} = -1$. Thus, by Claim 17 all other coefficients are 0 and the event has probability $p_i + p_j - 2p_i p_j = p_i(1 - p_j) + p_j(1 - p_i)$.

Let $c_{ij} = -1$. If $c_{kl} = 1$ for some pair k, l with $c_k = c_l = 0$, then the event has probability $p_i + p_j - p_ip_j - p_ip_k - p_lp_j + p_lp_k$, since by Claims 15, 16 and 17 all other coefficients must be 0. If $c_{kl} = 0$ for all k, l with $c_k = c_l = 0$, then by Claim 16 the event must have probability either $p_i + p_j - p_ip_j$, which is $1 - [(1 - p_i)(1 - p_j)]$, or $p_i + p_j - p_ip_j - p_ip_k$.

N=3: Let $c_i = c_j = c_k = 1$. In this case it can be seen that the cross coefficients $c_{ij} = c_{ik} = c_{jk} = -1$. Claims 17 and 16 imply that no coefficient $c_{j_1 j_2}$ can be nonzero, if $c_{j_1} = c_{j_2} = 0$. If any of the coefficients c_{il}, c_{jl} or c_{kl} is -1, for some $l \notin \{i, j, k\}$, then $\vec{v}\{i, j, k, l\} = -1$. Hence, all other coefficients are 0 and the event has probability $p_i + p_j + p_k - p_ip_j - p_jp_k - p_kp_i$. □

Note that we can define events which actually realize these possible probabilities. Let P, Q, R, S denote the pair-wise independent events $X_i = 1$, $X_j = 1$, $X_k = 1$, and $X_l = 1$, respectively. Then the listed probabilities in Theorem 13 are realized by the following events: $P\overline{P}$, P, PQ, $P\overline{Q}$, $\overline{P} \wedge \overline{Q}$, $(\overline{Q}P) \vee (QR)$, $(P\overline{Q}) \vee (Q\overline{P})$, $(P \wedge \overline{Q}) \vee (Q \wedge \overline{R}) \vee (R \wedge \overline{P})$, $(\overline{Q}P) \vee (Q\overline{R})$, $P[(RS\overline{Q}) \vee (\overline{R} \wedge \overline{(QS)})] \vee \overline{P}[SR \vee \overline{S}Q]$, $PQ \vee \overline{Q}R \vee (\overline{R} \wedge \overline{P})$.

If all the probabilities are equal to p then the only logical consequences have probabilities $0, p, p^2, p(1-p), 2p(1-p), 3p(1-p), (1-p)^2$ or their complements.

References

[1] N. Alon, L. Babai, and A. Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *Journal of Algorithms*, 7:567–583, 1986.

[2] N. Alon, J. Spencer, and P. Erdős. *The Probabilistic Method.* Wiley–Interscience Series, John Wiley & Sons, Inc., New York, 1992.

[3] B. Chor, O. Goldreich, J. Håstad, J. Friedman, S. Rudich, and R. Smolensky. The bit extraction problem or t–resilient functions. In *Proc. IEEE Symposium on Foundations of Computer Science*, pages 396–407, 1985.

[4] P. Frankl and R. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1:357–368, 1981.

[5] A. Joffe. On a set of almost deterministic k–independent random variables. *The Annals of Probability*, 2(1):161–162, 1974.

[6] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, IT–25(1):1–7, 1979.

[7] M. Luby. A simple parallel algorithm for the maximal independent set problem. *SIAM J. Comput.*, 15(4):1036–1053, 1986.

[8] L. J. Schulman. Small spaces uniform on neighborhoods. In *Proc. ACM Symposium on Theory of Computing*, pages 17–25, 1992.