On Construction of k-wise Independent Random Variables

Howard Karloff * Yishay Mansour †

Abstract

A *0-1 probability space* is a probability space $(\Omega, 2^\Omega, P)$, where the sample space $\Omega \subseteq \{0,1\}^n$ for some n. A probability space is *k-wise independent* if, when Y_i is defined to be the ith coordinate of the random n-vector, then any subset of k of the Y_i's is (mutually) independent, and it is said to be a probability space *for* $p_1, p_2, ..., p_n$ if $P[Y_i = 1] = p_i$.

We study constructions of k-wise independent 0-1 probability spaces in which the p_i's are arbitrary. It was known that for any $p_1, p_2, ..., p_n$, a k-wise independent probability space of size $m(n,k) = \binom{n}{k} + \binom{n}{k-1} + \binom{n}{k-2} + \cdots + \binom{n}{0}$ always exists. We prove that for some $p_1, p_2, ..., p_n \in [0,1]$, $m(n,k)$ is a lower bound on the size of any k-wise independent 0-1 probability space. For each fixed k, we prove that every k-wise independent 0-1 probability space for all $p_i = k/n$ has size $\Omega(n^k)$. For a very large degree of independence—$k = \lfloor \alpha n \rfloor$, for $\alpha > 1/2$—and all $p_i = 1/2$, we prove a lower bound on the size of $2^n(1 - \frac{1}{2\alpha})$. We also give explicit constructions of k-wise independent 0-1 probability spaces.

1 Introduction

A *0-1 probability space* is a probability space $(\Omega, 2^\Omega, P)$ whose sample space is a set of 0-1 strings of length n (for some n). It is *k-wise independent* if, when Y_i is defined to be the ith coordinate of the random n-vector, then any subset of k of the Y_i's is (mutually) independent, and it is said to be a probability space *for* $p_1, p_2, ..., p_n$ if $P[Y_i = 1] = p_i$.

A great deal of research in theoretical computer science recently has dealt with the problem of constructing small probability spaces satisfying certain independence constraints, most commonly k-wise independence [Joffe74, Lub86, ABI86, Lub88]. Aside from its interest as a fundamental issue, such small probability spaces can often be used for derandomizing randomized algorithms [ABI86, Lub86, KM93, Sch92, BR89, MNN89]: if one has a randomized algorithm which needs to be run on a k-wise independent 0-1 probability space, then it can be derandomized by running it deterministically on all the points in such a space (in parallel, if possible). This deterministic algorithm will be efficient if the size of the sample space is small, so that not too much time (or processors, in the parallel case) will be needed, and if the algorithm runs quickly on each point of the sample space.

*College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0280. This author was supported in part by NSF grant CCR 9107349.

†Computer Science Dept., Tel-Aviv University. This research was supported in part by the Israel Science Foundation administered by the Israel Academy of Science and Humanities and by a grant of the Israeli Ministry of Science and Technology.

Almost all previous work has dealt with the construction of *uniform* k-wise independent 0-1 probability spaces. These are spaces Ω in which all points $\omega \in \Omega$ have the same probability. For derandomization, nonuniform spaces are as good as uniform ones. A sample space Ω whose $|\Omega|$ points have many different probabilities will serve just as well. Suppose, for example, that a probabilistic construction needs a pairwise independent 0-1 probability space for $p_1, p_2, ..., p_n$. The conclusion of the existence proof is that some structure exists with positive probability. Then if we have a "small" pairwise independent 0-1 probability space Ω for $p_1, p_2, ..., p_n$, then we can try each of the points of Ω, and we will be guaranteed to find one for which the structure exists. The fact that the probabilities of the points in the sample space are not the same is irrelevant for this application. In our view, the construction of nonuniform k-wise independent 0-1 probability spaces has not received the appropriate attention.

We start with lower bounds. Define

$$m(n,k) = \binom{n}{k} + \binom{n}{k-1} + \binom{n}{k-2} + \cdots + \binom{n}{0}.$$

For constant k, $m(n,k)$ is $\Theta(n^k)$. The best general lower bound is from [ABI86, CGHFRS85, AS92].

Theorem 1. [ABI86, CGHFRS85, AS92] If $\mathcal{H}$ is a k-wise independent 0-1 probability space for $p_1, p_2, ..., p_n$ and all $p_i \in (0,1)$, then $|\mathcal{H}| \geq m(n, \lfloor k/2 \rfloor)$.

In fact, the theorem of [ABI86, CGHFRS85, AS92] is actually stronger, since it applies to random variables with arbitrary real range. We show that this lower bound is not always tight. We prove a lower bound of $m(n,k)$ for certain specified probabilities:

Theorem 2. Let n be a positive integer and let $1 \leq k \leq n$. Then there exist $p_1, p_2, ..., p_n \in (0,1)$ such that the size of every k-wise independent 0-1 probability space for $p_1, p_2, ..., p_n$ is at least $m(n,k)$.

Even when all the probabilities are identical, we can prove a lower bound on the size that is $\Omega(n^k)$, provided that k is fixed:

Theorem 3. Fix $k \geq 2$. The size of every k-wise independent 0-1 probability space if all $p_i = k/n$ is $\Omega(n^k)$.

Both Theorem 2 and Theorem 3 give much larger lower bounds than Theorem 1, since, if k is fixed in particular, then the lower bound of Theorem 1 is $O(n^{\lfloor k/2 \rfloor})$ while the bounds of Theorems 2 and 3 are $\Omega(n^k)$. Note that neither Theorem 2 nor Theorem 3 can be improved by allowing the probabilities to be arbitrary, for there *is* a k-wise independent 0-1 probability space if all $p_i = 1/2$ of size at most $2(2n)^{\lfloor k/2 \rfloor}$ [ABI86, AS92].

How close to optimal are these theorems? It is implicit in the work of D. Koller and N. Megiddo [KM93], who seem to have initiated the systematic study of nonuniform spaces in theoretical computer science, that there is always a space of size at most $m(n,k)$:

Theorem 4. [KM93] Let $n \geq 1$, let $1 \leq d \leq n$, and let $p_1, p_2, ..., p_n \in [0,1]$. Then there is a k-wise independent 0-1 probability space for $p_1, ..., p_n$ of size at most $m(n,k)$.

This implies that Theorem 2 is exactly tight and the bound of Theorem 3 is tight to within a constant factor (which depends on k but not n).

We also study k-wise independent 0-1 probability spaces when k is very large, say, $k = \lfloor \alpha n \rfloor$ for $\alpha \in (1/2, 1]$. In this case the best lower bound on size was $m(n, \lfloor k/2 \rfloor)$ (from Theorem 1). This quantity is bounded above by β^n, where $\beta < 2$ is a function only of α. Based on the techniques of [Fri92] we improve this bound as well.

Theorem 5. Let $1/2 < \alpha \leq 1$. Then the size of any $\lfloor \alpha n \rfloor$-wise independent 0-1 probability space for all $p_i = 1/2$ is at least $2^n(1 - \frac{1}{2\alpha})$.

Since it is trivial to build an n-wise independent 0-1 probability space for any p_i on $\Omega = \{0,1\}^n$, this theorem is also tight to within a constant factor. Furthermore, when $\alpha \in (0, 1/2)$, the upper bound of Theorem 4 implies that there is a k-wise independent 0-1 probability space of size at most γ^n for some $\gamma < 2$, so Theorem 5 cannot be extended to $\alpha < 1/2$.

We now turn to construction of small k-wise independent 0-1 probability spaces. We first give a general construction for arbitrary p_i. Namely, given $p_1, p_2, ..., p_n \in [0,1]$, which are possibly irrational, we give an explicit construction of a k-wise independent 0-1 probability space for $p_1, ..., p_n$. The size of the sample space is at most $(8n)^{2k}$. Note that for a fixed k the size of the sample space is bounded by a polynomial in the lower bound.

(In fact, the Koller-Megiddo technique shows how to generate in polynomial time a k-wise independent 0-1 probability space of size at most $m(n,k)$, for any $p_1, ..., p_n$, but it is not a "construction." One has to execute n times an algorithm that takes a feasible solution to a linear program and produces a basic feasible solution. On the other hand, their method is more general in that it is not restricted to k-wise independence.)

An interesting and well-studied special case of k-wise independent 0-1 random variables is the case in which $k = 2$ and all $p_i = p$. For this case we give a construction with a sample space of size $O(n^2)$, which is the best one could hope for if p is to be arbitrary, by Theorem 3. For some specific ranges we show that linear-size constructions are possible.

The paper is organized as follows. In Section 2 we have the lower bound in the case that the probabilities are different. Section 3 derives a lower bound in the case that the probabilities are the same. The lower bound for a large degree of independence appears in Section 4. The general construction with arbitrary probabilities is in Section 5. The construction for pairwise independent random variables is in Section 6.

2 Lower Bound for Different Probabilities

Throughout this section, let $0 < \gamma < 1/m(n,k)$ and let $p_i = \gamma^{2^{i-1}}$. We derive a tight lower bound of $m(n,k)$ on the size of any k-wise independent 0-1 probability space $\mathcal{H}$ for $p_1, p_2, ..., p_n$.

First we define a lexicographic ordering between sets.

Definition. Let $S = \{s_1, \ldots, s_l\}$ and $T = \{t_1, \ldots, t_j\}$ be distinct subsets of $\{1, 2, ..., n\}$, where $s_1 > s_2 > \cdots > s_l$ and $t_1 > t_2 > \cdots > t_j$. Then $S \succ T$ if either (1) T is a proper subset of S, or (2) there exists an m such that $s_i = t_i$ for $i = 1, 2, ..., m$ and $s_{m+1} > t_{m+1}$. Notice that either $S \succ T$ or $T \succ S$ if $S \neq T$.

For a set $S \subseteq \{1, 2, ..., n\}$, let $I_S(x)$ be the product of the values of x in coordinates $i \in S$, i.e., $I_S(x) = \prod_{i \in S} x_i$. Since the distribution is k-wise independent, for any specific set S of size at most k the probability of the event $[I_S(x) = 1]$ is exactly $\prod_{i \in S} p_i$, so let us use q_S to denote $\prod_{i \in S} p_i$. The following lemma shows that the q's of different sets are well separated.

Lemma 6. Let $S = \{s_1, \ldots, s_l\}$ and $T = \{t_1, \ldots, t_j\}$ be sets of size at most k such that $S \succ T$. Then

$$q_T/q_S = (p_{t_1} \cdots p_{t_j})/(p_{s_1} \cdots p_{s_l}) \geq 1/\gamma > m(n,k).$$

Proof. If T is a proper subset of S then the claim follows from the fact that for any i, $1/p_i \geq 1/\gamma > m(n,k)$.

Now assume T is not a proper subset of S. Without loss of generality $t_1 > t_2 > \cdots > t_j$, $s_1 > s_2 > \cdots > s_l$, $s_i = t_i$ for $i \leq m$, $s_{m+1} > t_{m+1}$.

$$p_{t_1} p_{t_2} \cdots p_{t_j} = (\prod_{i \leq m} p_{t_i})(\prod_{i > m} p_{t_i}).$$

$$p_{s_1} p_{s_2} \cdots p_{s_l} = (\prod_{i \leq m} p_{s_i})(\prod_{i > m} p_{s_i}).$$

Thus

$$q_T/q_S = \frac{\prod_{i>m} p_{t_i}}{\prod_{i>m} p_{s_i}}.$$

Let $r = t_{m+1}$. Notice that $s_{m+1} \geq t_{m+1} + 1 = r + 1$. The numerator of this quotient is at least

$$\prod_{i=1}^{r} p_i = \prod_{i=1}^{r} \gamma^{2^{i-1}} = \gamma^{2^r - 1}.$$

The denominator is at most

$$p_{s_{m+1}} = \gamma^{2^{s_{m+1}-1}} \leq \gamma^{2^r}.$$

Thus

$$q_T/q_S \geq \gamma^{(2^r - 1) - 2^r} = 1/\gamma > m(n,k). \blacksquare$$

Lemma 7. Let $\mathcal{H}$ be a k-wise independent 0-1 probability space for the $p_1, p_2, ..., p_n$ defined above. For any subset T of $\{1, 2, ..., n\}$ of size at most k, there is a vector x_T assigned positive probability by $\mathcal{H}$ such that $I_T(x_T) = 1$, and $I_S(x_T) = 0$ for all S such that $S \succ T$ and $|S| \leq k$.

Proof. Consider the following event:

$$[I_T(x) = 1 \text{ and } I_S(x) = 0 \text{ for all } S \text{ such that } S \succ T,\ |S| \leq k].$$

The probability of this event is bounded from below by

$$P[I_T(x) = 1] - \sum_{\substack{S \succ T, \\ |S| \leq k}} P[I_S(x) = 1].$$

Since the size of T is at most k and the size of each S is at most k, the probability of the event can be bounded from below via k-wise independence. The result is $q_T - \sum_{\substack{S \succ T, \\ |S| \leq k}} q_S$. By Lemma 6, for any $S \succ T$, $q_T \geq q_S/\gamma$. Since $m(n,k) < 1/\gamma$, we have $q_T - \sum_{\substack{S \succ T, \\ |S| \leq k}} q_S > 0$. Thus the probability of the event

$$[I_T(x) = 1 \text{ and } I_S(x) = 0 \text{ for all } S \text{ such that } |S| \leq k \text{ and } S \succ T]$$

is positive. It follows that there is some vector x_T assigned positive probability by $\mathcal{H}$ that has the required properties. ■

Now our theorem follows.

Theorem 8. Let B be any k-wise independent 0-1 probability space for the $p_1, ..., p_n$ defined above. Then $|B| \geq m(n, k)$.

Proof. By Lemma 7, for any set T of size at most k, there is a vector x_T having positive probability in $\mathcal{H}$ such that $I_T(x_T) = 1$ and also $I_S(x_T) = 0$ for any S of size at most k such that $S \succ T$. We need to show for any S and T, both of size at most k and satisfying $S \neq T$, that $x_S \neq x_T$. Without loss of generality assume that $S \succ T$. Hence, $I_S(x_T) = 0$. But by the definition of x_S, $I_S(x_S) = 1$. Therefore $x_S \neq x_T$. ■

Remark: It is easy to see that it is not necessary that each probability be exactly p_i; there is some slack in the proof. This means that for each n and i there is a subinterval of $[0, 1]$ of positive length so that if p'_i is in the given interval for all i, then the size of any k-wise independent 0-1 probability space for $p'_1, p'_2, ..., p'_n$ is at least $m(n, k)$.

3 Lower Bound for Identical Probabilities

In this section we study k-wise independent 0-1 probability spaces in which the p_i's all equal some p. In this scenario, for each fixed k, we are able to prove an $\Omega(n^k)$ lower bound on the size for $p = k/n$.

The proof is based on the following idea. Let $\mathcal{H}$ be a k-wise independent 0-1 probability space for all $p_i = p = k/n$. We consider any $S \subseteq \{1, 2, ..., n\}$ of size $k-1$. Since the distribution is k-wise independent, the probability of having 1's simultaneously in the positions in S is exactly p^{k-1}. Now we get simultaneous 1's in the positions of S by assigning some probability to the characteristic n-vector of S and by assigning the remaining probability to strings having 1's in all positions $i \in S$ and at least one additional 1. Informally, we define a $(k-1)$-set S to be *valid* if the sum of the probabilities assigned by $\mathcal{H}$ to all 0-1 strings having 1's in all positions $i \in S$ and at least one additional 1 is at least $p^{k-1}/3$. Note that since the distribution is k-wise independent, each string with k or more ones can be assigned probability at most p^k; so this means that if S is valid, the number of 0-1 strings in $\mathcal{H}$ with 1's in all positions $i \in S$ and at least one additional 1 is at least $[p^{k-1}/3]/p^k = 1/(3p)$. For $p = k/n$, $1/(3p) = n/(3k)$. This means that for each valid S, there are at least $n/(3k)$ 0-1 strings in $\mathcal{H}$ having 1's in all positions $i \in S$ and at least one additional 1. Furthermore, we show that for $p = k/n$, at least one quarter of the $(k-1)$-sets are valid. This would give us $\frac{1}{4}\binom{n}{k-1} \cdot \frac{n}{3k}$ strings in $\mathcal{H}$, if the strings obtained from valid set S were distinct from those obtained from valid set $T \neq S$. However, they need not be.

To circumvent this problem, we argue that for each valid S, there must actually be at least $n/(6k)^2$ strings in $\mathcal{H}$ having 1's in all positions in S, and no more than $6k-1$ additional 1's. (Obviously $n/(6k)^2$ is smaller than $n/(3k)$, but we gain in that the strings we examine have a limited number of 1's.) Now it is clear that any string with at most $(k-1)+(6k-1) = 7k-2$ 1's could have been generated from at most $\binom{7k-2}{k-1} < \binom{7k}{k-1}$ valid sets. This will give us a final tally of at least $[\frac{1}{4}\binom{n}{k-1} \cdot \frac{n}{(6k)^2}]/\binom{7k}{k-1}$ strings, and this is $\Omega(n^k)$ if k is fixed.

First, we need the following technical lemma. It will help us prove that there are many strings in $\mathcal{H}$

with few 1's apiece.

Lemma 9. Let $\alpha, \beta > 0$. Suppose $S_1, S_2, ..., S_m \subseteq \{1, 2, ..., N\}$ and $x_i > 0$ is associated with S_i such that

1. For $j = 1, 2, ..., N$, $\sum_{i:S_i \ni j} x_i \leq \alpha$, and
2. $\sum_{i=1}^{m} x_i \geq \beta$.

Then for any integer $l \geq 2$, there are at least

$$\frac{l\beta - N\alpha}{l(l-1)\alpha}$$

pairwise disjoint S_i's, each of size less than l.

Proof. Let $T_j = \{i : S_i \ni j\}$. From the assumptions about the x_i's, for $j = 1, 2, ..., N$,

$$\sum_{i \in T_j} x_i \leq \alpha.$$

Therefore

$$\sum_{j=1}^{N} \sum_{i \in T_j} x_i \leq N\alpha.$$

Each x_i is counted $|S_i|$ times on the left-hand side. Thus

$$\sum_{i=1}^{m} |S_i| x_i = \sum_{j=1}^{N} \sum_{i \in T_j} x_i \leq N\alpha.$$

Order the sets so that $|S_1| \leq |S_2| \leq \cdots \leq |S_m|$. Choose the minimal p so that $|S_p| \geq l$, choosing $p = m + 1$ if all $|S_i| < l$. Then

$$l \sum_{i=p}^{m} x_i \leq \sum_{i=p}^{m} |S_i| x_i \leq \sum_{i=1}^{m} |S_i| x_i \leq N\alpha$$

Therefore

$$\sum_{i=p}^{m} x_i \leq \frac{N\alpha}{l}.$$

Since the sum of all x_i's is bounded below by β,

$$\beta \leq \sum_{i=1}^{p-1} x_i + \sum_{i=p}^{m} x_i \leq \sum_{i=1}^{p-1} x_i + \frac{N\alpha}{l}.$$

Therefore

$$\sum_{i=1}^{p-1} x_i \geq \beta - \frac{N\alpha}{l}.$$

We may assume that $\beta - \frac{N\alpha}{l} > 0$, since otherwise the lemma is trivially true. Let $I = \emptyset$. Let $\mathcal{S} = \{1, 2, ..., p-1\}$. All sets S_i for $i \in \mathcal{S}$ have size less than l. Since $\sum_{i \in \mathcal{S}} x_i \geq \beta - \frac{N\alpha}{l} > 0$, there is a set S_t of size less than l with $t \in \mathcal{S}$.

Add t to I, and eliminate from $\mathcal{S}$ all i such that $S_i \cap S_t \neq \emptyset$ (including t itself). The sum of the x's of the sets remaining in $\mathcal{S}$ is at least

$$(\beta - \frac{N\alpha}{l}) - (l-1)\alpha,$$

since the sum of x_i over all S_i with S_i intersecting S_t is at most α times the size of S_t.

If $(\beta - \frac{N\alpha}{l}) - (l-1)\alpha > 0$, we repeat the process. This process can be repeated at least

$$\frac{\beta - \frac{N\alpha}{l}}{(l-1)\alpha}$$

times. At the end, we have a set of at least

$$\frac{\beta l - N\alpha}{l(l-1)\alpha}$$

pairwise disjoint sets, each of size less than l. ■

Let $k \geq 2$. Let $0 < p \leq 1$. Suppose that $\mathcal{H}$ is a k-wise independent 0-1 probability space for all $p_i = p$.

Definition. In a k-wise independent 0-1 probability space $\mathcal{H}$, say a $(k-1)$-subset S of $\{1, 2, ..., n\}$ is *valid* if for an x drawn randomly from $\mathcal{H}$,

$$P[(I_S(x) = 1) \text{ and } (x \text{ has } k \text{ or more 1's})] > \frac{1}{3}p^{k-1}.$$

We first show that at least one quarter of the sets of size $k-1$ are valid sets.

Lemma 10. For every k-wise independent 0-1 probability space $\mathcal{H}$ such that $p_i = p = k/n$ for $1 \leq i \leq n$, at least one quarter of the sets of size $k-1$ are valid.

Proof. If S is not valid, then the characteristic vector of S must be assigned probability exceeding $p^{k-1} - \frac{1}{3}p^{k-1} = \frac{2}{3}p^{k-1}$ by $\mathcal{H}$. Clearly, the sum of the probabilities assigned by $\mathcal{H}$ to characteristic vectors of $(k-1)$-sets is at most 1. Therefore the number of invalid $(k-1)$-sets is at most $(3/2)/p^{k-1}$. We will show that for $p = k/n$ this is at most $(3/4)\binom{n}{k-1}$, and therefore at least one quarter of the sets of size $k-1$ are valid.

Let $p = k/n$. We have

$$\frac{1}{p^{k-1}} = \left(\frac{n}{k}\right)^{k-1} \leq \frac{1}{2}\binom{n}{k-1},$$

using $\binom{a}{b} \geq (a/b)^b$ if $a \geq b \geq 1$ and $[k/(k-1)]^{k-1} \geq 2$ for all $k \geq 2$. This means that the number of invalid $(k-1)$-sets is at most

$$\frac{3/2}{p^{k-1}} \leq \frac{3}{4}\binom{n}{k-1}$$

and therefore the number of valid sets is at least

$$\binom{n}{k-1} - \frac{3}{4}\binom{n}{k-1} = \frac{1}{4}\binom{n}{k-1}. \blacksquare$$

Theorem 11. Every k-wise independent 0-1 probability space $\mathcal{H}$ for all $p_i = p = k/n$ has size at least

$$\frac{\frac{1}{4}\binom{n}{k-1}\frac{n}{(6k)^2}}{\binom{7k}{k-1}}.$$

If $n \geq 2k$, this is at least

$$c \cdot \frac{m(n,k)}{k^3 2^{7k}}$$

for a universal constant $c > 0$.

Proof. Identify $B \subseteq \{1, 2, ..., n\}$ with its characteristic n-vector. We use $P[B]$ to denote the probability assigned by $\mathcal{H}$ to the characteristic vector of B.

Suppose $\{1, 2, ..., k-1\}$ is valid. Let $U_1, U_2, ..., U_m$ be the proper supersets of $\{1, 2, ..., k-1\}$ such that $P[U_i] > 0$. Let $x_i = P[U_i]$ and $S_i = U_i - \{1, 2, ..., k-1\}$. For $j \geq k$, let $T_j = \{i : S_i \ni j\}$. Because $\mathcal{H}$ is a k-wise independent 0-1 probability space for all $p_i = p$, the chance of having 1's simultaneously in positions $1, 2, 3, ..., k-1$ and j is exactly p^k. But this probability is exactly $\sum_{i \in T_j} x_i$, so this is p^k. We have

1. $S_i \subseteq \{k, k+1, ..., n\}$.
2. For $j = k, k+1, ..., n$, $\sum_{i \in T_j} x_i = p^k$, and
3. $\sum_{i=1}^{m} x_i > (1/3)p^{k-1}$, because $\{1, 2, ..., k-1\}$ is valid.

Apply Lemma 9 with $N = n - (k-1)$, $\alpha = p^k$, $\beta = (1/3)p^{k-1}$. By the lemma, for all $l \geq 2$, there are at least

$$\frac{\frac{1}{3}lp^{k-1} - (n-k+1)p^k}{l(l-1)p^k} = \frac{\frac{1}{3p} - \frac{(n-k+1)}{l}}{(l-1)}$$

pairwise disjoint S_i's, each of size at most $l-1$. Now set $l = 6k$ and use $p = k/n$. We have at least $\frac{n}{(6k)^2}$ pairwise disjoint S_i's, each of size at most $6k-1$. Thus there are at least $\frac{n}{(6k)^2}$ strings having positive probability in $\mathcal{H}$ with 1's in positions $1, 2, ..., k-1$ and at most $(k-1) + (l-1) = 7k-2$ 1's in total.

Now this argument obviously goes through for each of the $\frac{1}{4}\binom{n}{k-1}$ or more valid sets. This means that, *counting multiplicities*, there are at least $\frac{1}{4}\binom{n}{k-1} \cdot \frac{n}{(6k)^2}$ strings with at most $7k-2$ 1's each. Since each such string contains at most $\binom{7k-2}{k-1} < \binom{7k}{k-1}$ subsets of size $k-1$, each such string is generated fewer than $\binom{7k}{k-1}$ times by different valid sets. Thus the number of *different* strings of positive probability exceeds

$$\frac{\frac{1}{4}\binom{n}{k-1}\frac{n}{(6k)^2}}{\binom{7k}{k-1}}.$$

Proving that this last quantity is at least $c \cdot \frac{m(n,k)}{k^3 2^{7k}}$ for a universal constant $c > 0$ if $k \leq n/2$ is not difficult, if one uses facts such as $(k+1)\binom{n}{k} \geq m(n,k)$ if $k \leq n/2$, $\binom{n}{k-1}n \geq \binom{n}{k}$, and $\binom{7k}{k-1} \leq 2^{7k}$. ■

Remark: Again, it is not necessary that p_i be *exactly* k/n. For each n there is a subinterval I of $[0,1]$ of positive length so that if $p \in I$ and each $p_i = p$, then every k-wise independent 0-1 probability space for all $p_i = p$ is of size $\Omega(n^k)$, for fixed k.

4 Lower Bounds for Large Independence

In this section we show how to derive a lower bound for the case in which the degree of independence is very large. The main tool in proving our result is the following theorem, in which $\mathcal{H}_d$ denotes the set of all d-dimensional subcubes of the n-dimensional hypercube $\mathcal{H} = \{1, -1\}^n$.

Theorem 12. Let $n \geq 1$. Let g be a 2^n-dimensional probability vector viewed as a function on the vertices of the n-dimensional hypercube $\mathcal{H}$. Let $1 \leq d \leq n$. Let $N = |\{v \in \mathcal{H} : g(v) > 0\}|$. Suppose that there is an α such that for all $H \in \mathcal{H}_d$, $\alpha = \sum_{v \in H} g(v)$. Then $N \geq 2^{n-1}(n+2-2d)/(n+1-d)$.

The proof of the above theorem uses the techniques of [Fri92]. Before giving the proof, we derive an interesting consequence.

Corollary 13. Let B be a k-wise independent 0-1 probability space for all $p_i = 1/2$. Then $|B| \geq 2^n(\frac{2k+2-n}{2k+2}) = 2^n(1 - \frac{n}{2k+2})$.

Proof. For all $(n-k)$-dimensional subcubes H, $\sum_{v \in H} g(v) = 2^{-k}$. (We are using here the fact that all $p_i = 1/2$.) So let $d = n - k$, $\alpha = 2^{-k}$. Then $N \geq 2^{n-1}(n+2-2(n-k))/(n+1-(n-k)) = 2^n(2k-n+2)/(2+2k)$. ■

Corollary 14. Let $0 < \alpha \leq 1$ and $\lfloor \alpha n \rfloor \leq k \leq n$, and suppose that B is a k-wise independent 0-1 probability space for all $p_i = 1/2$. Then

$$|B| \geq \left(1 - \frac{1}{2\alpha}\right) 2^n.$$

Proof. Because $k \geq \lfloor \alpha n \rfloor$, $1 - \frac{n}{2(k+1)} \geq 1 - \frac{1}{2\alpha}$. By Corollary 13, we have $|B| \geq 2^n(1 - \frac{n}{2k+2}) \geq 2^n(1 - \frac{1}{2\alpha})$. ■

Now we do the proof of Theorem 12. We start with the following technical lemma.

Lemma 15. Let g be a probability vector. Suppose that $N = |\{i | g(i) > 0\}|$. Then $|g|^2 \geq 1/N$.

Proof. By convexity of x^2, $|g|^2$ is minimized when all the probability is pushed onto as many points as possible. ■

We use A to denote the $2^n \times 2^n$ adjacency matrix of $\mathcal{H}$. For each $T \subseteq \{1, 2, ..., n\}$, let x_T be a vector, viewed as a function on $\mathcal{H}$, whose vth coordinate $x_T(v) = \prod_{i \in T} v_i$. If $f, h : \mathcal{H} \to \mathbb{R}$, let $< f, h >= 2^{-n} \sum_{x \in \mathcal{H}} f(x)h(x)$. This is the normal dot product, scaled by 2^{-n}.

Lemma 16. Let $0 \leq r \leq n$. Then $n - 2r$ is an eigenvalue of A, and for each $T \subseteq \{1, 2, ..., n\}$ of size r, x_T is an eigenvector of A corresponding to eigenvalue $n - 2r$.

Proof. Let $T \subseteq \{1, 2, ..., n\}$ with $|T| = r$. It suffices to prove that

$$Ax_T = (n - 2r)x_T.$$

Choose any $u \in \mathcal{H}$. The (u, v) entry $A(u, v)$ of A is 1 if u and v differ in exactly one position, and 0 otherwise. This means that the uth entry of Ax_T is

$$\sum_{v \in H} A(u, v)x_T(v) = \sum_{v \in N(u)} x_T(v),$$

where $N(u)$ denotes those $v \in \mathcal{H}$ differing from u in exactly one position. Of the n v's in $N(u)$, $n-r$ v's differ from u in a bit position not in T. These v's satisfy $x_T(v) = x_T(u)$. The remaining r v's differ from u in one position in T and satisfy $x_T(v) = -x_T(u)$. It follows that

$$\begin{aligned} \textstyle\sum_{v \in N(u)} x_T(v) &= (n-r)x_T(u) - r x_T(u) \\ &= (n-2r)x_T(u). \end{aligned}$$

Since u was arbitrary, we have

$$Ax_T = (n-2r)x_T. \blacksquare$$

It is easy to see that if $T \neq U$, then $< x_T, x_U >= 0$, and that $< x_T, x_T >= 1$ for all T. This means that the x_T's are orthogonal and hence $\{x_T | T \subseteq \{1,2,...,n\}\}$ is linearly independent. Thus any 2^n-dimensional vector g can be written as a linear combination $g = \sum_T \beta_T x_T$ for some reals β_T.

Proof of Theorem 12. As the theorem is vacuous if $d = n$, we assume without loss of generality that $d < n$.

Fixing g, let us set $\alpha_U =< g, x_U >$ for any U. It is easy to see, from the orthogonality of the x_T's, that $g = \sum_T \alpha_T x_T$. Let us use g_r to denote $\sum_{|T|=r} \alpha_T x_T$; clearly $g = \sum_{r=0}^n g_r$.

We are given that

$$\sum_{v \in H} g(v) = \alpha$$

for all $H \in \mathcal{H}_d$. Choose any $U \subseteq \{1,2,...,n\}$ such that $1 \leq |U| \leq n-d$.

$$\begin{aligned} \alpha_U &= < g, x_U > \\ &= 2^{-n} \textstyle\sum_v g(v) x_U(v) \end{aligned}$$

Let us use "$v|_U = s$" to mean that the restriction of v to the bit positions in U is s.

$$\begin{aligned} 2^n \alpha_U &= \textstyle\sum_v g(v) x_U(v) \\ &= \textstyle\sum_{s \in \{1,-1\}^{|U|}} \sum_{v:v|_U = s} g(v) x_U(v) \\ &= \textstyle\sum_{s \in \{1,-1\}^{|U|}} (\prod_{i=1}^{|U|} s_i)(\sum_{v:v|_U=s} g(v)). \end{aligned}$$

Now $n - |U| \geq d$ and $\{v : v|_U = s\}$ is a cube of dimension $n - |U|$. This $(n-|U|)$-dimensional cube can be partitioned into $\frac{2^{n-|U|}}{2^d}$ d-dimensional subcubes. As the sum of $g(v)$ over any d-dimensional subcube is α,

$$\sum_{v:v|_U=s} g(v) = \frac{2^{n-|U|}}{2^d}\alpha.$$

Thus

$$\sum_{s \in \{1,-1\}^{|U|}} (\prod_{i=1}^{|U|} s_i)(\sum_{v:v|_U=s} g(v)) = \left(\frac{2^{n-|U|}}{2^d}\alpha\right) \sum_{s \in \{1,-1\}^{|U|}} \prod_{i=1}^{|U|} s_i.$$

Now

$$\sum_{s \in \{1,-1\}^{|U|}} \prod_{i=1}^{|U|} s_i = [1 + (-1)]^{|U|} = 0,$$

since $|U| \geq 1$. Therefore $2^n \alpha_U = 0$ and $\alpha_U = 0$. Since $g_r = \sum_{|T|=r} \alpha_T x_T$, it follows that $g_r = 0$ if $1 \leq r \leq n-d$.

Now let $\mu_r = |g_r|^2/|g|^2$ for $0 \le r \le n$. Notice that $\mu_r = 0$ if $1 \le r \le n-d$.

$$\sum_{r=0}^{n} \mu_r = \sum_{r=0}^{n} \frac{|g_r|^2}{|g|^2}.$$

Now $g = \sum g_r$ and $< g_r, g_s >= 0$ if $r \ne s$ imply that $|g|^2 = \sum_{r=0}^n |g_r|^2$. Therefore $\sum_{r=0}^n \mu_r = 1$.

$$\begin{aligned} < Ag, g > &= < A\textstyle\sum_{r=0}^n g_r, \sum_{s=0}^n g_s > \\ &= < \textstyle\sum_{r=0}^n Ag_r, \sum_{s=0}^n g_s > \\ &= \textstyle\sum_{r=0}^n \sum_{s=0}^n < Ag_r, g_s > . \end{aligned}$$

Using $g_r = \sum_{|T|=r} \alpha_T x_T$, we have, for arbitrary r, s,

$$\begin{aligned} < Ag_r, g_s > &= < \textstyle\sum_{|T|=r} \alpha_T A x_T, \sum_{|U|=s} \alpha_U x_U > \\ &= \textstyle\sum_{|T|=r, |U|=s} \alpha_T \alpha_U < Ax_T, x_U > . \end{aligned}$$

For $|T| = r$, $n - 2r$ is an eigenvalue corresponding to x_T. Therefore $Ax_T = (n-2r)x_T$. Hence

$$\begin{aligned} < Ag_r, g_s > &= \textstyle\sum_{|T|=r, |U|=s} \alpha_T \alpha_U < (n-2r)x_T, x_U > \\ &= \textstyle\sum_{|T|=r, |U|=s} \alpha_T \alpha_U (n-2r) < x_T, x_U > . \end{aligned}$$

If $r \ne s$, this is 0. For $r = s$, we have

$$\begin{aligned} < Ag_r, g_r > &= \textstyle\sum_{|T|=r} \alpha_T^2 (n-2r) \\ &= (n-2r) \textstyle\sum_{|T|=r} \alpha_T^2. \end{aligned}$$

Now

$$\begin{aligned} |g_r|^2 &= < g_r, g_r > \\ &= < \textstyle\sum_{|T|=r} \alpha_T x_T, \sum_{|U|=r} \alpha_U x_U > \\ &= \textstyle\sum_{|T|=r} \alpha_T^2. \end{aligned}$$

Hence $< Ag_r, g_r >= (n-2r)|g_r|^2$, and if $r \ne s$, then $< Ag_r, g_s >= 0$. Thus

$$< Ag, g >= \sum_{r=0}^{n} \sum_{s=0}^{n} < Ag_r, g_s >= \sum_{r=0}^{n} < Ag_r, g_r >= \sum_{r=0}^{n} (n-2r)|g_r|^2.$$

Thus

$$< Ag, g >= |g|^2 \sum_{r=0}^{n} (n-2r)\mu_r.$$

Clearly $< Ag, g > \ge 0$. Now

$$0 \le < Ag, g >= |g|^2 [(n - 2\cdot 0)\mu_0 + 0 + \sum_{r=n-d+1}^{n} (n-2r)\mu_r].$$

So

$$\begin{aligned} 0 &\le n\mu_0 + \textstyle\sum_{r=n-d+1}^n (n-2r)\mu_r \\ &\le n\mu_0 + (2d-2-n) \textstyle\sum_{r=n-d+1}^n \mu_r. \end{aligned}$$

So

$$\begin{aligned} 0 &\le n\mu_0 + (2d-2-n)(1-\mu_0). \\ 0 &\le \mu_0(n - 2d + 2 + n) + (2d - 2 - n) \\ n + 2 - 2d &\le 2\mu_0(n + 1 - d) \\ \mu_0 &\ge \tfrac{1}{2}\left(\tfrac{n+2-2d}{n+1-d}\right). \end{aligned}$$

Now $g_0 = \alpha_\emptyset x_\emptyset$ with $\alpha_\emptyset =< x_\emptyset, g >= 2^{-n}e^T g$, with e being the all-1's vector. This means that $|g_0|^2 = \alpha_\emptyset^2 |x_\emptyset|^2 = 2^{-2n}(e^T g)^2 2^n = 2^{-n}(e^T g)^2$. Thus

$$|g_0|^2 = \frac{1}{2^n}(e^T g)^2.$$

So

$$\mu_0 = \frac{|g_0|^2}{|g|^2} = \frac{1}{2^n}\frac{(e^T g)^2}{|g|^2} = \frac{1}{2^n}\frac{1}{|g|^2}.$$

So

$$\frac{1}{|g|^2} = 2^n \mu_0 \geq \frac{2^n}{2}\left(\frac{n+2-2d}{n+1-d}\right).$$

By Lemma 15, $|g|^2 \geq \frac{1}{N}$. Thus $\frac{1}{|g|^2} \leq N$. It follows that

$$N \geq \frac{1}{|g|^2} \geq \frac{2^n}{2}\left(\frac{n+2-2d}{n+1-d}\right). \blacksquare$$

5 General Construction

In this section we give a general construction for k-wise independent 0-1 random variables. The construction receives as input the desired $p_1, p_2, ..., p_n$. There is no limitation on the p_i; they do not even have to be rational numbers (for example, $p_i = 1/\pi^i$). As output we give a set of vectors B, where for each $x \in B$ we associate a probability. The distribution is simply to choose an $x \in B$, according to the given probabilities.

We first need to state some lemmas. As before, we use q_S to denote $\prod_{i \in S} p_i$.

Lemma 17. Let $\mathcal{H}$ be a 0-1 probability space. Define Y_i to be the ith bit of the random n-vector. Let $1 \leq k \leq n$. Suppose that there are reals $p_1, p_2, ..., p_n$ such that for all subsets S of $\{1, 2, ..., n\}$ of size at most k,

$$P[\cap_{i \in S}[Y_i = 1]] = q_S.$$

Then $\mathcal{H}$ is a k-wise independent 0-1 probability space for $p_1, p_2, ..., p_n$.

In other words, it suffices to ensure that the probability that any k or fewer variables are simultaneously 1 is what it should be. There is no need to prove explicitly, for example, that $P[Y_2 = 0, Y_4 = 1, Y_5 = 0] = (1-p_2)p_4(1-p_5)$.

Proof. This is an easy consequence of the Moebius inversion formula (actually, of inclusion-exclusion). Fix $U \subseteq \{1, 2, ..., n\}$, $|U| = k$. We prove that the variables $\{Y_i | i \in U\}$ are independent, by proving that for all $T \subseteq U$, the probability that $Y_i = 1$ for all $i \in T$ and $Y_i = 0$ for all $i \in U - T$ is exactly

$$\left(\prod_{i \in T} p_i\right)\left(\prod_{i \in U-T}(1-p_i)\right).$$

For any subset S of U, let

$$f(S) = P[\cap_{i \in S}[Y_i = 1]]$$

and let

$$g(S) = P[(\cap_{i \in S}[Y_i = 1]) \cap (\cap_{i \in U-S}[Y_i = 0])].$$

(We define $f(S)$ and $g(S)$ only for subsets of U.) Clearly

$$f(S) = \sum_{T:T\supseteq S} g(T)$$

for any S. We now apply the Moebius inversion formula. It implies that, given the $f(S)$'s, the $g(T)$'s are determined uniquely. (Actually, the Moebius inversion formula applies if we are summing over subsets, not supersets, but it applies here to f', g' where $f'(S') = f(U - S')$, $g'(T') = g(U - T')$.) But we are given that $f(S) = q_S$, and we know that one solution to the system

$$f(S) = \sum_{T:T\supseteq S} g(T) \;\; \forall S \subseteq U$$

is

$$g(T) = (\prod_{i\in T} p_i)(\prod_{i\in U-T}(1-p_i)).$$

So this must be the unique solution. Since U was an arbitrary subset of k variables, $Y_1, Y_2, ..., Y_n$ are k-wise independent. ■

Here is another simple lemma:

Lemma 18. Suppose B is a k-wise independent 0-1 probability space for $p_1, p_2, ..., p_n$. Interchanging 0's and 1's in the first bit position of every point in the probability space gives a k-wise independent 0-1 probability space for $1 - p_1, p_2, p_3, ..., p_n$.

The proof is easy and is omitted.

Corollary 19. Suppose B is a k-wise independent 0-1 probability space for $p_1, p_2, ..., p_n$. Suppose that $p'_i = 1 - p_i$ if $i \in S$, $p'_i = p_i$ otherwise. A k-wise independent 0-1 probability space for $p'_1, p'_2, ..., p'_n$ can be obtained from B by interchanging 0's and 1's in every position $i \in S$ in each point of the probability space.

Proof. Follows easily from Lemma 18. ■

Our general construction for $p_1, p_2, ..., p_n$ consists of three parts: (1) a construction for probabilities very close to 1, (2) a construction for $p'_1, p'_2, ..., p'_n$ that are approximations to the desired probabilities $p_1, p_2, ..., p_n$, and (3) a method of combining the two probability spaces so as to get $p_1, ..., p_n$ exactly.

First, we must give a construction for very large probabilities. Since it is easier to understand what's going on when the probabilities are very small, we will assume $p_i \leq 1/(2n)$ for all i, and then invoke Corollary 19 to get the space for $p_i \geq 1 - \frac{1}{2n}$.

Theorem 20. Let $\epsilon = 1/(2n)$. Suppose that $p_1, p_2, ..., p_n \leq \epsilon$. There is a k-wise independent 0-1 probability space $\mathcal{H}$ for $p_1, p_2, ..., p_n$ whose points are the elements of $\{0,1\}^n$ with at most k 1's. Its size is at most $m(n, k)$.

Proof. We assign a probability to each 0-1 vector as follows. Identify $B \subseteq \{1, 2, ..., n\}$ with its characteristic n-vector. We use $P[B]$ to denote the probability that $\mathcal{H}$ assigns to the characteristic vector of B. Set $P[B] = 0$ if $|B| > k$. We will create probabilities $P[B]$ for each $B \subseteq \{1, 2, ..., n\}$, $|B| \leq k$, such that for all $A \subseteq \{1, 2, ..., n\}$, $|A| \leq k$,

$$\sum_{B:B\supseteq A} P[B] = q_A = \prod_{i\in A} p_i.$$

Where Y_i denotes the random variable which is the ith bit of the random n-vector, this means that in $\mathcal{H}$,

$$P[\cap_{i\in A}[Y_i = 1]] = \sum_{B:B\supseteq A} P[B] = q_A.$$

By Lemma 17, this is enough.

We assign probabilities $P[B]$ by downward induction on $|B|$. As each $P[B]$ is defined, we prove inductively that $0 \le P[B] \le q_B$. If $|B| = k$, set $P[B] = q_B$. (Notice that $0 \le P[B] \le q_B$.)

Now suppose that $|B| = l$, $l < k$, and $P[C]$ has been defined for all C of size exceeding l. Set

$$P[B] = q_B - \sum_{C:C\supsetneq B} P[C] = q_B - \sum_{C:C\supsetneq B, |C|\le k} P[C].$$

It is obvious with this definition that

$$q_B = \sum_{C:C\supseteq B} P[C].$$

Now if $0 \le P[C] \le q_C$ for all C of size exceeding l, then clearly $P[B] \le q_B$. Also,

$$\begin{aligned} P[B] &\ge q_B - \textstyle\sum_{C:C\supsetneq B,|C|\le k} q_C \\ &\ge q_B - q_B \textstyle\sum_{i=1}^{k-l} \binom{n-l}{i} \epsilon^i \\ &\ge q_B[1 - \textstyle\sum_{i=1}^{\infty} (n\epsilon)^i] \\ &= q_B[1 - \textstyle\sum_{i=1}^{\infty} (\frac{1}{2})^i] \\ &= 0. \end{aligned}$$

That $m(n,k)$ is the number of 0-1 vectors with at most k 1's is obvious. ∎

Interchanging 0's and 1's in the construction of Theorem 20, we get a k-wise independent 0-1 probability space for arbitrary $p_i \ge 1 - \frac{1}{2n}$.

Now we state a well-known lemma [Joffe74].

Lemma 21. Let $r \ge n$ be a prime. Choose $a_0, a_1, ..., a_{k-1}$ uniformly and independently at random from $\mathbb{Z}_r$, and let $X_i = a_0 + a_1 i + a_2 i^2 + \cdots + a_{k-1} i^{k-1} \bmod r$, for $1 \le i \le r$. Then $X_1, X_2, ..., X_r$ are k-wise independent and each X_i is uniform on $\mathbb{Z}_r$. Furthermore, suppose $p_1, p_2, ..., p_r \in [0,1]$ satisfy $p_i = j_i/r$ for all i for integers j_i. Define a probability space with (at most) r^k points, as follows. For each point $< a_0, a_1, ..., a_{k-1} >$ in the sample space, let $Y_i = 1$ if $X_i \in \{0, 1, ..., j_i - 1\}$ and $Y_i = 0$ otherwise, $1 \le i \le r$. Associate with $< a_0, a_1, ..., a_{k-1} >$ the vector $< Y_1, Y_2, ..., Y_r >$ and give each of the r^k r-vectors probability $1/r^k$ (and then combine identical r-vectors, if necessary). Then the new 0-1 probability space defined by the vectors $< Y_1, Y_2, ..., Y_r >$ is a k-wise independent 0-1 probability space for $p_1, p_2, ..., p_r$.

To recap, we now have: (1) a construction for arbitrary $p_i \ge 1 - \frac{1}{2n}$ and (2) a construction for probabilities j_i/r for $i = 1, 2, ..., n$, where $r \ge n$ is a prime. We must show how to combine these two spaces to get a new one.

We give a construction that takes a k-wise independent 0-1 probability space A for $p_1, p_2, ..., p_n$, and a second k-wise independent 0-1 probability space A' for $p'_1, p'_2, ..., p'_n$, and produces a k-wise independent 0-1 probability space B for $p_1p'_1, p_2p'_2, ..., p_np'_n$. To generate an element of B, independently pick x from A and x' from A', and output the bit-wise AND $x \wedge x' \in \{0,1\}^n$. Equivalently, B can be obtained by taking each x in the sample space of A and each x' in the sample space of A' and adding $x \wedge x'$ to the sample space of B, with probability equal to the product of the probabilities in A and in A'. (Then remove duplicates by summing their probabilities.) Clearly $|B| \leq |A||A'|$.

Lemma 22. The above construction of B from A and A' guarantees that B is a k-wise independent 0-1 probability space for $p_1p'_1, p_2p'_2, ..., p_np'_n$.

Proof. From Lemma 17, it is sufficient to show that for any subset $S \subseteq \{1, 2, ..., n\}$ of size at most k, $P[\cap_{i \in S}[Y_i = 1]] = \prod_{i \in S}(p_i p'_i)$, where Y_i is the ith bit of the random n-vector $x \wedge x'$. But the ith bit of $x \wedge x'$ is 1 if and only if the ith bit of x and the ith bit of x' are both one. By independence, this occurs with probability

$$(\prod_{i \in S} p_i)(\prod_{i \in S} p'_i). \blacksquare$$

Based on the above lemma we exhibit the general construction. First note that we can assume $p_i \geq 1/2$. Otherwise we can set p_i to $1 - p_i$, and at the end of the construction switch between 0 and 1 for for ith variable.

We create a probability distribution that is k-wise independent such that the probability of each single variable is between the desired probability and the desired probability plus $1/(4n)$. Let p_i be the desired probability, and r a prime in the range $[4n, 8n]$. Then first we create a probability space A using Lemma 21 with $j_i = \lceil p_i r \rceil$ for all i. This guarantees that the expectation s_i of each single variable is in the desired range (and is $s_i = \lceil p_i r \rceil / r$). Let this probability space be A, of size at most $r^k \leq (8n)^k$.

Now we need to "adjust" the distribution. Let $t_i = p_i / s_i \leq 1$. Because $r > 4n$,

$$1 - \frac{1}{2n} < 1 - \frac{2}{r}.$$

Let $z_i = rp_i \geq \frac{1}{2}r$. Now

$$t_i = \frac{p_i}{s_i} = \frac{z_i}{\lceil z_i \rceil} = 1 - \frac{\lceil z_i \rceil - z_i}{\lceil z_i \rceil} \geq 1 - \frac{1}{\lceil z_i \rceil}$$
$$\geq 1 - \frac{1}{r/2} = 1 - \frac{2}{r}.$$

Thus

$$1 - \frac{1}{2n} < 1 - \frac{2}{r} \leq t_i \leq 1.$$

Therefore we can create a k-wise independent 0-1 probability space for $t_1, t_2, ..., t_n$, by Theorem 20, of size at most $m(n, k)$. Let us call this probability space A'.

The last step is to merge the two constructions. The merging is done through Lemma 22. The expectation of the ith variable will be $t_i s_i = p_i$. The size of the new space is at most $(8n)^k m(n, k)$.

Theorem 23. The new space B is a k-wise independent 0-1 probability space for any sequence of probabilities $p_1, p_2, ..., p_n$. Its size is at most $(8n)^k m(n, k) \leq (8n)^{2k}$.

6 Construction for Pairwise Independence

In this section we exhibit a space of size at most $4n^2$ for pairwise independent random variables in the case that all the probabilities are equal. From the lower bound of Theorem 11 it follows that this construction is optimal in size, to within a constant factor.

Definition. Let $\mathcal{H}$ be a 0-1 probability space. $\mathcal{H}$ is an *$(a, b; n)$-space* if the probability of a 1 in position i is a for all i, and the probability of simultaneous 1's in positions i and j is b, for all $i < j$.

For a pairwise independent 0-1 probability space, we would want $b = a^2$. However, here we will need $b < a^2$.

The following lemma shows that given such an $(a, b; n)$-space, we can build a *pairwise independent* 0-1 probability space if all $p_i = p$, provided that p is in a certain interval:

Lemma 24. Let B be an $(a, b; n)$-space with $b < a^2$. Then there is a construction of a pairwise independent 0-1 probability space for all $p_i = p$, on a sample space of size at most $|B| + 2$, provided that $p \in [b/a, (a - b)/(1 - a)]$.

Proof. The new probability space will be defined in the following way. Let $x, y \geq 0$ such that $x + y \leq 1$. With probability x, choose a point in B (according to the probabilities in B). With probability y, choose the all 1's row. With probability $1 - x - y$, choose the all-0's row. It is clear that the implied probability space has size at most $|B| + 2$.

In order that the new space be pairwise independent for all $p_i = p$, it suffices to ensure that the probability of a 1 in position i is p and the probability of simultaneous 1's in positions i and j is p^2 (and that the sum of the probabilities of the points is 1). In other words, it is sufficient to satisfy the following system:

$$\begin{aligned} x \cdot a + y \cdot 1 &= p \\ x \cdot b + y \cdot 1 &= p^2 \end{aligned}$$

and satisfy $x + y \leq 1$, $x, y \geq 0$. The unique solution to the system is

$$\begin{aligned} x &= (p - p^2)/(a - b), \\ y &= p(ap - b)/(a - b). \end{aligned}$$

We have $x + y \leq 1$, $x, y \geq 0$ if and only if

$$p \in [b/a, (a - b)/(1 - a)].$$

Note that since $b < a^2$, $b/a < (a - b)/(1 - a)$. ■

Based on the above lemma we can exhibit a quadratic size space for an arbitrary probability p.

Theorem 25. For any $n > 1$, for any $0 \leq p \leq 1$, there is a construction of a pairwise independent 0-1 probability space of size at most $4n^2$ if all $p_i = p$.

Proof. First we claim that if $q \geq n$ is a prime, then for any l, $0 \leq l \leq q$, there is an $(l/q, (l^2 - l)/(q^2 - q); n)$-space. The points of the sample space are the $q^2 - q$ pairs (c, d), $c, d \in \mathbb{Z}_q$, $d \neq 0$. Let $X_i(c, d) = c + di \bmod q$ and $Y_i(c, d) = 1$ if $X_i(c, d) \in \{0, 1, 2, ..., l - 1\}$ and $Y_i(c, d) = 0$ otherwise, $1 \leq i \leq n$. Build a space with at most $q^2 - q$ vectors by associating with point (c, d) the n-vector $(Y_1, Y_2, ..., Y_n)$, and then combining identical n-vectors.

That this is an $(l/q, (l^2 - l)/(q^2 - q); n)$-space can easily be verified by noting that the q omitted pairs $(c, 0)$ define l rows that are all 1 and $q - l$ rows that are all 0, and if the q omitted rows were added, we'd have a pairwise independent space in which each random variable is 1 with probability l/q.

Now for the construction of the pairwise independent probability space. Choose a prime q such that $n \leq q \leq 2n$. Choose an l, $1 \leq l \leq q-1$, such that $p \in [(l-1)/(q-1), l/(q-1)]$. The construction of this proof gives us an $(l/q, (l^2-l)/(q^2-q); n)$-space of size at most q^2-q. Lemma 24 now gives us a pairwise independent sample space, of size at most $(q^2-q)+2 \leq q^2 \leq (2n)^2 = 4n^2$, with $p_i = p$ for all i, provided that $p \in [b/a, (a-b)/(1-a)]$. A simple calculation with $a = l/q$, $b = (l^2-l)/(q^2-q)$ shows that $[b/a, (a-b)/(1-a)] = [(l-1)/(q-1), l/(q-1)]$, which contains p. ■

For certain ranges of p, better constructions are possible. Theorem 26 uses Lemma 24 to get a linear construction for $p \leq 1/(n-1)$, while Corollary 29 exhibits a linear construction for values in the neighborhood of $1/2$.

Theorem 26. For $p \leq 1/(n-1)$ there is a pairwise independent 0-1 probability space of size at most $n+2$ with $p_i = p$ for all i.

Proof. We exhibit a $(1/n, 0; n)$-space; this is enough for all $p \in [b/a, (a-b)/(1-a)] = [0, (1/n)/(1-1/n)] = [0, 1/(n-1)]$. This is simply the identity matrix in which each row gets probability $1/n$. By Lemma 24 there is a construction of size at most $n+2$. ■

We can also get linear size constructions around $1/2$.

Lemma 27. Let H be a normalized $m \times m$ Hadamard matrix (i.e., an $m \times m$ $+1, -1$ matrix satisfying $H^T H = mI$, whose first row and first column are all $+1$'s). Let H' be obtained from H by omitting the first row and first column and then replacing all -1's by 0's. The probability space obtained by giving each of the $m-1$ rows of H' probability $1/(m-1)$ is a $((1/2)(m-2)/(m-1), (1/4)(m-4)/(m-1); m-1)$-space.

The proof is simple and is omitted.

Theorem 28. Let $n > 1$ and let H be a normalized $m \times m$ Hadamard matrix H with $m \geq n+1$. There is a pairwise independent 0-1 probability space (whose points are elements of $\{0,1\}^n$) of size at most $m+1$ with $p_i = p$ for all i if $p \in [1/2 - 1/(m-2), 1/2 + 1/(m-2)]$.

Proof. With $a = (1/2)(m-2)/(m-1)$, $b = (1/4)(m-4)/(m-1)$ (from Lemma 27), we have $b/a = 1/2 - 1/(m-2)$ and $(a-b)/(1-a) = 1/2$. This gives a construction for $p \in [1/2 - 1/(m-2), 1/2]$. Its size is at most $(m-1)+2 = m+1$. For $p \in [1/2, 1/2 + 1/(m-2)]$, we first build a space for $1-p$ and then exchange 0's and 1's. ■

Corollary 29. For any $n > 1$, for any $p \in [1/2 - 1/(2n-2), 1/2 + 1/(2n-2)]$, there is a pairwise independent 0-1 probability space of size at most $2n+1$ with all $p_i = p$.

Proof. For any t, there is a $2^t \times 2^t$ Hadamard matrix. Given n, take m to be the least power of two exceeding n; $m \leq 2n$. By Theorem 28, there is a construction of size at most $m+1 \leq 2n+1$ for any $p \in [1/2 - 1/(m-2), 1/2 + 1/(m-2)] \supseteq [1/2 - 1/(2n-2), 1/2 + 1/(2n-2)]$. ■

7 Acknowledgments

For their help in the early stages of our work, we are very grateful to D. Koller and N. Megiddo, whose paper [KM93], in opening up the area of nonuniform spaces in theoretical computer science, inspired this one. We thank Shi-Chun Tsai for pointing out a typo.

References

[ABI86] N. Alon, L. Babai, and A. Itai, "A Fast and Simple Randomized Parallel Algorithm for the Maximal Independent Set Problem," *Journal of Algorithms* 7 (1986), pages 567–583.

[AS92] N. Alon and J. Spencer, *The Probabilistic Method,* Wiley, 1992.

[BR89] B. Berger and J. Rompel, "Simulating ($log^c n$)-wise Independence in NC," *Proc.* 30^{th} *IEEE Symposium on Foundations of Computer Science*, 1989, pages 2–7.

[CGHFRS85] B. Chor, O. Goldreich, J. Hastad, J. Friedman, S. Rudich, and R. Smolensky, "The Bit Extraction Problem or t-Resilient Functions," *Proc.* 26^{th} *IEEE Symposium on Foundations of Computer Science*, 1985, pages 396–407.

[Fri92] J. Friedman, "On the Bit Extraction Problem," *Proc. 33rd IEEE Symposium on Foundations of Computer Science*, 1992, pages 314–319.

[Joffe74] A. Joffe, "On a Set of Almost Deterministic k-Independent Random Variables," *Annals of Probability* 2 (1974), pages 161–162.

[KM93] D. Koller and N. Megiddo, Constructing Small Sample Spaces Satisfying Given Constraints," *Proc.* 25^{th} *ACM Symposium on Theory of Computing*, 1993, pages 268–277.

[Lub86] M. Luby, "A Simple Parallel Algorithm for the Maximal Independent Set Problem," *SIAM Journal on Computing* 15 (1986), pages 1036–1053.

[Lub88] M. Luby, "Removing Randomness in Parallel Computation Without a Processor Penalty," *Proc.* 29^{th} *IEEE Symposium on Foundations of Computer Science*, 1988, pages 162-173.

[MNN89] R. Motwani, J. Naor, and J. Naor, "The Probabilistic Method Yields Deterministic Parallel Algorithms," *Proc.* 30^{th} *IEEE Symposium on Foundations of Computer Science*, 1989, pages 8–13.

[Sch92] L. Schulman, "Sample Spaces Uniform on Neighborhoods," *Proc.* 24^{th} *ACM Symposium on Theory of Computing*, 1992, pages 17–25.