On the joint entropy of d-wise-independent variables

Dmitry Gavinsky[1 2], Pavel Pudlák[1]

Abstract. How low can the joint entropy of n d-wise independent (for $d \geq 2$) discrete random variables be, subject to given constraints on the individual distributions (say, no value may be taken by a variable with probability greater than p, for $p < 1$)? This question has been posed and partially answered in a recent work of Babai [*Entropy versus pairwise independence* (preliminary version), `http://people.cs.uchicago.edu/ laci/papers/13augEntropy.pdf`, 2013].

In this paper we improve some of his bounds, prove new bounds in a wider range of parameters and show matching upper bounds in some special cases. In particular, we prove tight lower bounds for the min-entropy (as well as the entropy) of pairwise and three-wise independent balanced binary variables for infinitely many values of n.

Keywords: d-wise-independent variables; entropy; lower bound

Classification: 60C05

1. Introduction

Suitable choice of a (discrete) distribution is a crucial component that underlies many results in extremal combinatorics and theoretical computer sciences (e.g., see [AS08]). It is often the case that the "ideal" distribution to use would be mutually independent over n random variables $X_1 \ldots, X_n$ (each variable taking one of several possible values); however, "full" mutual independence is "too expensive" and a d-wise-independent distribution is used instead (e.g., see [LW06]). (A string of random variables $X_1, \ldots, X_n$ is called d-wise independent if any d-tuple of the variables is independent.) Indeed, if all variables are independent, then the sample space has at least exponential size, while d-wise independent spaces can be of polynomial size if d is constant. This has many applications in computer science. The size of the space, the number of random bits needed and the joint entropy of $X_1, \ldots, X_n$ are closely related parameters that are crucial in these applications.

This is a motivation of the question studied in a recent article of Babai [Bab13]: what is the minimum entropy for n pairwise independent variables. Babai showed

DOI 10.14712/1213-7243.2015.169

[1]Partially funded by the grant P202/12/G061 of GAČR and by RVO: 67985840.

[2]Part of this work was done while visiting the Centre for Quantum Technologies at the National University of Singapore, and was partially funded by the Singapore Ministry of Education and the NRF.

an asymptotically logarithmic lower bound, by proving a very nice theorem. He proved that for any string $X_1, \ldots, X_n$ of pairwise independent binary-valued variables, where the probabilities are bounded away from zero and one, there exists a logarithmic size subset of these variables that is almost independent. Such a subset must have entropy asymptotically equal to its size, so the logarithmic lower bound follows.

Our aim in this paper is to answer some questions and improve bounds of Babai. For proving tight bounds, a more traditional approach (see for example [Lan65]) based on a construction of orthogonal matrices seems more suitable. This approach enables us, first, to extend Babai's bounds to a larger range of parameters, and second, to obtain more precise bounds. In particular, we prove that the joint entropy of $X_1, \ldots, X_n$ is logarithmic even if the entropy of the variables is only of the order of $\log n/n$, which is the lowest possible. Furthermore, we prove a lower bound $\log(n+1)$, conjectured by Babai, on the min-entropy of pairwise independent balanced binary variables (i.e., when each X_j is equal to 0, respectively 1, with probability 1/2). This matches the upper bounds given by the well known construction based on Hadamard matrices. So the bound is tight if an Hadamard matrix of dimension $n+1$ exists.

Lower bounds on the entropy of d-wise independent variables can be obtained from lower bounds on pairwise independent variables by a well-known construction that produces a longer string of pairwise independent variables. We slightly modify this construction for odd values of d which enables us to obtain a matching upper and lower bounds for $d = 3$ and infinitely many values of n (powers of 2).

Although we are primarily interested in binary-valued variables, we will show that some of our lower bounds can be extended to the case of general (finite-outcome) pairwise independent variables.

2. Preliminaries

We will write

$$H[X] = \sum_x \mathbf{Pr}[X = x] \cdot \log \frac{1}{\mathbf{Pr}[X = x]}$$

to denote the Shannon entropy of the (discrete) random variable X, and

$$H_{min}[X] = \min_x \log \frac{1}{\mathbf{Pr}[X = x]}$$

for the min-entropy. Clearly, $H_{min}[X] \le H[X]$. All logarithms are to the base 2. Random variables $X_1 \ldots, X_n$ are said to be d-wise independent if for every $s \in \binom{[n]}{d}$, the variables $(X_i)_{i \in s}$ are mutually independent. A random variable is called *binary* if it is supported on $\{0, 1\}$.

Recall Cantelli's inequality [Can10] — a strengthening of Chebyshev's inequality for the case of one-sided deviations:

Lemma 2.1 (Cantelli's inequality)**.** *For every random variable X and real $t > 0$,*

$$\mathbf{Pr}\left[X \leq \mathbf{E}[X] - t\right],\ \mathbf{Pr}\left[X \geq \mathbf{E}[X] + t\right] \leq \frac{1}{1 + \frac{t^2}{\mathbf{Var}\left[X\right]}}. \tag{1}$$

3. Lower bounds

In this section we give lower bounds on the joint entropy of n d-wise-independent variables.

3.1 Pairwise independent binary variables. Here we give two incomparable entropy lower bounds for the families of pairwise independent binary variables.

Theorem 3.1. *Let $X = (X_1 \dots, X_n)$ be n pairwise independent binary variables. Let $q_j = \mathbf{Pr}\left[X_j = 1\right]$ and suppose that $0 < q_j < 1$ for $j = 1 \dots, n$. Then*

$$H[X] \geq \sup_{0<t\leq n} \frac{\log(n + 1 - t)}{1 + \frac{1}{t^2} \sum_{j=1}^n \frac{(1-2q_j)^2}{q_j(1-q_j)}}.$$

Proof: Let $A = \{a_1 \dots, a_m\} \subseteq \{0,1\}^n$ be the support of X and $p_i \stackrel{\text{def}}{=} \mathbf{Pr}\left[X = a_i\right]$. We will denote by a_{ij} the j'th element of a_i for $j \in [n]$.

Define an $m \times (n+1)$ matrix $U = \{u_{ij}\}$ as follows. For all $i \in [m]$,

$$u_{i0} \stackrel{\text{def}}{=} \sqrt{p_i},$$

and for $j \in [n]$,

$$u_{ij} \stackrel{\text{def}}{=} \begin{cases} -\sqrt{\frac{p_i q_j}{1-q_j}} & \text{if } \ a_{ij} = 0, \\ \sqrt{\frac{p_i(1-q_j)}{q_j}} & \text{if } \ a_{ij} = 1. \end{cases}$$

For $0 \leq j \leq n$, let u_j denote the j'th column vector of U; note that these vectors form an orthonormal family: For $j > 0$,

$$\langle u_0, u_j \rangle = \sqrt{\frac{1-q_j}{q_j}} \cdot \mathbf{Pr}\left[X_j = 1\right] - \sqrt{\frac{q_j}{1-q_j}} \cdot \mathbf{Pr}\left[X_j = 0\right] = 0;$$

for $k > j > 0$,

$$\begin{aligned}\langle u_k, u_j\rangle = &\sqrt{\frac{1-q_k}{q_k}}\cdot\left(\sqrt{\frac{1-q_j}{q_j}}\cdot\mathbf{Pr}\left[X_k=1\wedge X_j=1\right]\right.\\ &\left.-\sqrt{\frac{q_j}{1-q_j}}\cdot\mathbf{Pr}\left[X_k=1\wedge X_j=0\right]\right)\\ &+\sqrt{\frac{q_k}{1-q_k}}\cdot\left(\sqrt{\frac{1-q_j}{q_j}}\cdot\mathbf{Pr}\left[X_k=0\wedge X_j=1\right]\right.\\ &\left.-\sqrt{\frac{q_j}{1-q_j}}\cdot\mathbf{Pr}\left[X_k=0\wedge X_j=0\right]\right)\\ = &\,0,\end{aligned}$$

as follows from independence of X_i and X_j. As well, the norm of every u_i is 1.

Since the matrix U is unitary, or can be made unitary by adding more columns, we know that the norm of each row of U is at most 1. Thus we get, for all $i \in [m]$,

$$\begin{aligned}1 \geq \sum_{j=0}^{n} u_{ij}^2 &= p_i\cdot\left(1+\sum_{j:a_{ij}=0}\frac{q_j}{1-q_j}+\sum_{j:a_{ij}=1}\frac{1-q_j}{q_j}\right)\\ &= p_i\cdot\left(1+\sum_{j=0}^{n}\left((1-a_{ij})\frac{q_j}{1-q_j}+a_{ij}\frac{1-q_j}{q_j}\right)\right).\end{aligned}\tag{2}$$

Our aim is to find a subset $A_0 \subseteq A$ such that every string $a \in A_0$ has a low probability, whereas the weight of A is large. This will give us the lower bound on the entropy.

Let A_0 be all the elements of A satisfying

$$1+\sum_{j=0}^{n}\left((1-a_{ij})\frac{q_j}{1-q_j}+a_{ij}\frac{1-q_j}{q_j}\right) \geq n+1-t.\tag{3}$$

Without loss of generality we may assume that $A_0 = \{a_1 \ldots, a_{m_0}\}$. Then, according to (2), for every $i = 1, \ldots, m_0$,

$$p_i \leq 1/(n+1-t).\tag{4}$$

Let Y be the random variable defined by

$$Y := 1+\sum_{j=1}^{n}\left((1-X_j)\frac{q_j}{1-q_j}+X_j\frac{1-q_j}{q_j}\right) = 1+\sum_{j=1}^{n}\frac{q_j}{1-q_j}+\sum_{j=1}^{n}\frac{1-2q_j}{q_j(1-q_j)}\cdot X_j.$$

Then we have

$$\sum_{i=1}^{m_0} p_j = \mathbf{Pr}\,[Y \geq n+1-t].$$

The expectation of Y is

$$\mathbf{E}[Y] = 1 + \sum_{j=1}^{n} \left((1-q_j)\frac{q_j}{1-q_j} + q_j \frac{1-q_j}{q_j} \right) = n+1.$$

The variance of Y is

$$\begin{aligned}\mathbf{Var}\,[Y] &= \mathbf{Var}\left[\sum_{j=1}^{n} \frac{1-2q_j}{q_j(1-q_j)} \cdot X_j\right] = \sum_{j=1}^{n} \mathbf{Var}\left[\frac{1-2q_j}{q_j(1-q_j)} \cdot X_j\right] \\ &= \sum_{j=1}^{n} \left(\frac{1-2q_j}{q_j(1-q_j)}\right)^2 \cdot \mathbf{Var}\,[X_j] = \sum_{j=1}^{n} \frac{(1-2q_j)^2}{q_j(1-q_j)},\end{aligned}$$

where we have used the fact that the variables are pairwise independent. Now we apply Cantelli's inequality (Lemma 2.1) to the random variable Y and parameter t.

$$\begin{aligned}\sum_{i=1}^{m_0} p_j &= \mathbf{Pr}\,[Y \geq n+1-t] = \mathbf{Pr}\,[Y \geq \mathbf{E}[Y] - t] \\ &\geq \frac{1}{1+\frac{1}{t^2}\,\mathbf{Var}\,[Y]} = \frac{1}{1+\frac{1}{t^2}\sum_{j=1}^{n}\frac{(1-2q_j)^2}{q_j(1-q_j)}}.\end{aligned}$$

Using this inequality and the fact that $p_i^{-1} > n+1-t$ for all $i \in [m_0]$ (which is (4)), we get

$$H[X] = \sum_{i=1}^{m} p_i \log p_i^{-1} \geq \sum_{i=1}^{m_0} p_i \log p_i^{-1} > \frac{\log(n+1-t)}{1+\frac{1}{t^2}\sum_{j=1}^{n}\frac{(1-2q_j)^2}{q_j(1-q_j)}},$$

as required. □

Suppose that $0 < q \leq q_j \leq 1/2$ for some q and all j. Since $\frac{1}{q} \geq \frac{1-2q_j}{q_j(1-q_j)}$, we have

$$H[X] \geq \sup_{0<t\leq n} \frac{\log(n+1-t)}{1+\frac{n}{t^2 q}}. \tag{5}$$

In particular, for $t = n/2$,

$$H[X] \geq \frac{\log(n/2+1)}{1+\frac{4}{nq}}.$$

This proves that the entropy of X is $\Omega(\log n)$ as long as $q_j \geq \epsilon n^{-1}$, $j = 1, \ldots, n$, for some $\epsilon > 0$, i.e., if $H[X_j] = \Omega(\log n/n)$. On the other hand, if $q_j \leq q(n)$, $j = 1, \ldots, n$, for some $q(n) = o(n^{-1})$, then $H[X_j] = o(n^{-1}\log n)$, and thus $H[X] = o(\log n)$.

If all $q_j = 1/2$ we get $H[X] \geq \log(n+1)$ by taking $t \to 0$. This is tight for infinitely many values of n (see Section 4) and confirms Conjecture 1.2 of Babai [Bab13]. However, the following theorem implies the same bound even for the min-entropy and the proof is, in fact, more direct.

Theorem 3.2. *Let $X = (X_1 \ldots, X_n)$ be n pairwise independent binary variables. Let $q_j = \mathbf{Pr}[X_j = 1]$ and suppose that $0 < q_j < 1$ for $j = 1 \ldots, n$. Then*

$$H_{min}[X] \geq \log\left(1 + \sum_{j=1}^{n} \min\left\{\frac{1-q_j}{q_j}, \frac{q_j}{1-q_j}\right\}\right).$$

Proof: Let U be an $m \times (n+1)$ matrix as in the proof of Theorem 3.1, assuming again without loss of generality that $\mathbf{Pr}[X_j = 1] \leq \mathbf{Pr}[X_j = 0]$ always. From (2) we get that for all $i \in [m]$,

$$1 \geq p_i \cdot \left(1 + \sum_{j:a_{ij}=0} \frac{q_j}{1-q_j} + \sum_{j:a_{ij}=1} \frac{1-q_j}{q_j}\right) \geq p_i \cdot \left(1 + \sum_{j=1}^{n} \min\left\{\frac{1-q_j}{q_j}, \frac{q_j}{1-q_j}\right\}\right),$$

which gives us the required lower bound on p_i's. □

Corollary 3.3. *If all $q_j \geq q$, then*

$$H_{min}[X] \geq \log\left(1 + \frac{nq}{1-q}\right). \qquad (6)$$

For $q = 1/2$ (unbiased X_i's), this corollary gives

$$H_{min}[X] \geq \log(n+1),$$

which is tight for infinitely many values of n.

3.2 Pairwise independent finite-outcome variables. Let $[k]$ be the values that a random variable X_j takes on, $k \geq 2$.

Theorem 3.4. *Let $X = (X_1 \ldots, X_n)$ be pairwise independent variables that take on values in $[k]$, $k \geq 2$. Let w be such that for all $i \in [n]$, $j \in [k]$,*

$$\mathbf{Pr}[X_i = j] \leq w \text{ (i.e., } H_{min}[X_i] \geq -\log w).$$

If $w \geq 1/2$, then

$$H_{min}[X] \geq \log\left(\frac{1-w}{w} \cdot n + 1\right).$$

If $w \leq 1/2$, then

$$H_{min}[X] \geq \log(n+1).$$

To prove the theorem, we need the following technical statement.

Claim 3.5. For $k \geq 2$, let $b_1 \dots, b_k \geq 0$ be such that $\sum_{t=2}^{k} b_t \geq b_1$ and for all $t \geq 2$, $b_t \leq b_1$. Then there exist $\alpha_2 \dots, \alpha_k \in \mathbb{R}$, such that

$$\sum_{t=2}^{t} \mathrm{e}^{\mathrm{i}\alpha_t} b_t = b_1.$$

Proof of Claim 3.5: Let C_r denote the circle in the complex plane with radius r and center in 0. The claim is equivalent to the statement that b_1 is in the Minkowski sum of $C_{b_2}, \dots, C_{b_k}$. Note that if $r \leq s$, then $C_r + C_s$ contains C_s as a subset. Thus the sum $C_2 + \dots + C_k$ is either a region between $C_{b_2+\dots+b_k}$ and some smaller circle, or a disc with radius $b_2 + \dots + b_k$ — in any case, it contains both $C_{\max_{2\leq t\leq k} b_t}$ and $C_{b_2+\dots+b_k}$. Hence it also contains b_1. □

Proof of Theorem 3.4: The proof is a modification of the proofs of Theorems 3.1 and 3.2.

Let $A = \{a_1 \dots, a_m\} \subseteq [k]^n$ be the support of X and $p_i \stackrel{\text{def}}{=} \mathbf{Pr}\left[X = a_i\right]$. For $j \in [n]$, let $w_j \stackrel{\text{def}}{=} \max\{\mathbf{Pr}\left[X_j = t\right] \mid t \in [k]\}$ and assume without loss of generality that $\mathbf{Pr}\left[X_j = 1\right] = w_j$. Let $\omega_j = \max\{1, \frac{w_j}{1-w_j}\}$, and let $\alpha_{j2} \dots, \alpha_{jk}$ be the values guaranteed by Claim 3.5 for $b_1 = w_j/\omega_j$ and $b_t = \mathbf{Pr}\left[X_j = t\right]$ for $2 \leq t \leq k$ (which observes the claim requirements).

This time we define the matrix U over $\mathbb{C}$: for $i \in [m]$,

$$u_{i0} \stackrel{\text{def}}{=} \sqrt{p_i},$$

and for $j \in [n]$,

$$u_{ij} \stackrel{\text{def}}{=} \begin{cases} -\sqrt{p_i/\omega_j} & \text{if } \; a_{ij} = 1, \\ \sqrt{p_i\omega_j} \cdot \mathrm{e}^{\mathrm{i}\alpha_{jz}} & \text{if } \; a_{ij} = z > 1. \end{cases}$$

As before, let u_j denote the j'th column vector of U. Then, by the immediate adaptation of the argument we gave for Theorem 3.1 (taking into account the guarantees of Claim 3.5), it holds that for all $j \neq k > 0$,

$$\langle u_j, u_k \rangle = 0 \quad \text{and} \quad \|u_j\| = 1.$$

Therefore, the norm of each row of U is at most 1 and, for every i,

$$1 \geq p_i \cdot \left(1 + \frac{n}{\omega_{max}}\right),$$

where $\omega_{max} \stackrel{\text{def}}{=} \max\{\omega_j \mid j \in [n]\}$. The result follows. □

3.3 d-wise independent unbiased binary variables. One can use an idea from [ABI86] to derive from the case of pairwise independent variables stronger lower bounds for d-wise independent variables. We will demonstrate it only on Theorem 3.2, but the same idea can be used together with other lower bounds on the entropy of pairwise independent variables.

Theorem 3.6. *Let $X = (X_1 \dots, X_n)$ where X_j's are d-wise independent unbiased binary variables. If d is even, then*

$$H_{min}[X] \geq \log\left(\sum_{i=0}^{d/2}\binom{n}{i}\right).$$

If d is odd, then

$$H_{min}[X] \geq \log\left(\sum_{i=0}^{(d-1)/2}\binom{n}{i} + \binom{n-1}{(d-1)/2}\right).$$

Proof: Let d be even. We define $Y = (Y_1 \dots, Y_m)$, where all Y_i's are unbiased binary variables equal to the parity of at most $d/2$ variables X_i and $m = \sum_{i=1}^{d/2}\binom{n}{i}$ (every Y_i is unique). Clearly, $Y_1 \dots, Y_m$ are pairwise independent, and from Theorem 3.2 we get

$$H_{min}[X] \geq H_{min}[Y] \geq \log\left(1 + \sum_{i=1}^{d/2}\binom{n}{i}\right) = \log\left(\sum_{i=0}^{d/2}\binom{n}{i}\right).$$

If d is odd, we take the parities of at most $(d-1)/2$ variables X_i and the parities of X_1 with exactly $(d-1)/2$ other variables. The resulting variables are again pairwise independent. $\square$

In the next section we will see, in particular, that the above bound is tight for the case of $d = 3$ and n being a power of 2.

4. Upper bounds

In this section we review some constructions of d-wise independent unbiased binary variables with low entropies. The constructions are based on known ideas, and they are included here to argue optimality of the lower bounds from Section 3.

The standard way of constructing d-wise independent distributions is using parity check matrices of codes with minimum distance at least d. In such matrices every d columns are linearly independent. Hence, if we take the space of vectors generated by the rows of such a matrix, i.e., the dual code, we obtain d-wise independent variables. Over GF_2, these are balanced binary variables. To get matching bounds we have to find suitable codes.

We start with the case of pairwise independent variables ($d = 2$). Recall that an Hadamard matrix is a real matrix with entries ± 1 whose rows (and hence also columns) are orthogonal. Hadamard matrices exist for infinitely many dimensions, in particular for every power of 2. Given an Hadamard matrix of dimension $n+1$, first transform it into an Hadamard matrix with the first column having all 1s, then delete the first column. The resulting $(n+1)\times n$ matrix defines in a natural way an *Hadamard code* and n pairwise independent balanced binary variables supported on a set of size $n+1$.

Lancaster [Lan65] proved:

1. For every $n \geq 2$, there exist at most n pairwise independent random variables on a probability space with $n+1$ points.
2. The existence of such random variables where, additionally, each point in the probability space has measure $\frac{1}{n+1}$ is equivalent to the existence of an Hadamard matrix of dimension $n+1$.

Our proofs of Theorems 3.1 and 3.2 can be viewed as an extension of an argument used by Lancaster to prove 2. Lancaster considered general (finite-outcome) pairwise independent variables. For unbiased binary variables, we can prove the following.

Theorem 4.1. *The existence of n pairwise independent unbiased binary variables with entropy equal to* $\log(n+1)$ *is equivalent to the existence of an Hadamard matrix of dimension* $n+1$.

Proof: As shown above, an Hadamard matrix of dimension $n+1$ gives rise to n pairwise independent unbiased binary variables with entropy equal to $\log(n+1)$.

To prove the converse, let n pairwise independent unbiased binary variables with entropy equal to $\log(n+1)$ be given. According to Theorem 3.2, every point in the probability space has measure at most $\frac{1}{n+1}$. Since the entropy is $\log(n+1)$, this implies that there are exactly $n+1$ points, each with measure $\frac{1}{n+1}$. The existence of an Hadamard matrix of dimension $n+1$ then follows from Lancaster's theorem, or from our proof of Theorem 3.1. $\square$

Another case where we can precisely match the lower bound for infinitely many values of n is $d=3$. Let $n=2^l$ and consider the $(l+1)\times n$ binary matrix whose first row consists of 1's and the columns restricted to the remaining l rows are all vectors of length l. Every two different columns are linearly independent over $\mathcal{G}F_2^{l+1}$ because they are different. Every three different columns are also independent because every two of them are and they cannot sum to zero vector due to the first row. Hence the space generated by the rows is 3-wise independent. The size of the space is $2^{l+1}=2n$, precisely matching the statement of Theorem 3.6. Thus we have:

Theorem 4.2. *If n is a power of 2, then the minimum of $H_{min}[X]$ taken over all n-tuples of 3-wise independent unbiased binary variables is* $\log 2n$.

Note that the above construction is based on the parity-check matrix of the Hamming code: first we extend the matrix by a column with all zeros and then we extend it by a row with all ones. The two constructions, one based on the

Hadamard code and the other based on the Hamming code, can be generalized using BCH codes. Recall that the binary BCH code of length 2^m-1 and designed distance $2t+1$ has the minimal distance at least $2t+1$ and dimension 2^m-1-mt, provided that m is sufficiently large with respect to t (see [MS83], pages 258 and 253). Hence every $2t$ columns of the parity-check matrix (and also of the dual code) are linearly independent and the dimension of the space generated by the parity-check matrix (i.e., the dual code) has dimension mt. Thus for $d>2$ even, we can take a BCH code with designed distance $2t+1=d+1$ and we get $n=2^m-1$ d-wise independent random variables with min-entropy

$$\frac{d}{2}\log(n+1).$$

For $d>3$ odd, we take $2t+1=d$, and extend the parity-check matrix by a column of zeros and a row of ones, as we did above. Thus we obtain a matrix with every d columns independent. Let $n=2^m$ be the number of columns of this matrix. The linear space generated by the rows gives a probability space of n d-wise independent random variables with min-entropy

$$\frac{d-1}{2}\log n+1.$$

These bounds are asymptotically equal to the lower bounds of Theorem 3.6 when n goes to infinity. However, we have not been able to find constructions matching our lower bound exactly for any $d\geq 4$ and any n.

5. Conclusions

We proved several lower bounds on the entropy of pairwise and d-wise independent random variables. Our lower bounds match upper bounds exactly, or asymptotically for some special values of the parameters involved. But for most values of parameters, we do not know even the asymptotic behavior of the dependence of entropy on them. This is, in particular, so in the case of equally distributed pairwise independent 0-1 variables. In this special case we have two bounds (5) and (6), which give an asymptotically optimal bound for $q\approx 1/n$ and a tight bound for $q=1/2$, but for other values we do not know. Another interesting problem, studied in [Bab13], is to find the best lower bound on the joint entropy $H[X_1,\dots,X_n]$ of a string of pairwise random variables $X_1,\dots,X_n$ in terms of the parameter $L:=\sum_j X_j$. For more open problems, see [Bab13].

Acknowledgment. We would like to thank an anonymous referee for suggesting better formulations of some theorems and an elegant proof of Claim 3.5.

References

[ABI86] Alon N., Babai L., Itai A., *A fast and simple randomized parallel algorithm for the maximal independent set problem*, J. Algorithms **7** (1986), no. 4, 567–583.

[AS08] Alon N., Spencer J., *The Probabilistic Method*, John Wiley, Hoboken, NJ, 2008.

[Bab13] Babai L., *Entropy versus pairwise independence* (preliminary version), `http://people.cs.uchicago.edu/ laci/papers/13augEntropy.pdf`, 2013.

[Can10] Cantelli F.P., *Intorno ad un teorema fondamentale della teoria del rischio*, Bollettino dell' Associazione degli Attuari Italiani **24** (1910), 1–23.

[Lan65] Lancaster H.O., *Pairwise statistical independence*, Ann. Math. Statist. **36** (1965), no. 4, 1313–1317.

[LW06] Luby M., Wigderson A., *Pairwise independence and derandomization*, Found. Trends Theor. Comput. Sci. **1** (2005), no. 4, 237–301.

[MS83] MacWilliams F.J., Sloane N.J.A., *The Theory of Error-Correcting Codes*, North Holland Publishing Co., Amsterdam-New York-Oxford, 1977.

Institute of Mathematics, Academy of Sciences, Žitná 25, Praha 1, Czech Republic

(*Received* April 10, 2016)