SHARP ESTIMATES FOR GOWERS NORMS ON DISCRETE CUBES

TONĆI CRMARIĆ AND VJEKOSLAV KOVAČ

Abstract. We study optimal dimensionless inequalities

$$\|f\|_{\mathrm{U}^k} \leqslant \|f\|_{\ell^{p_{k,n}}}$$

that hold for all functions $f\colon \mathbb{Z}^d \to \mathbb{C}$ supported in $\{0,1,\ldots,n-1\}^d$ and estimates

$$\|\mathbb{1}_A\|_{\mathrm{U}^k}^{2^k} \leqslant |A|^{t_{k,n}}$$

that hold for all subsets A of the same discrete cubes. A general theory, analogous to the work of de Dios Pont, Greenfeld, Ivanisvili, and Madrid, is developed to show that the critical exponents are related by $p_{k,n}t_{k,n} = 2^k$. This is used to prove the three main results of the paper:

- an explicit formula for $t_{k,2}$, which generalizes a theorem by Kane and Tao,
- two-sided asymptotic estimates for $t_{k,n}$ as $n \to \infty$ for a fixed $k \geqslant 2$, which generalize a theorem by Shao, and
- a precise asymptotic formula for $t_{k,n}$ as $k \to \infty$ for a fixed $n \geqslant 2$.

1. Introduction

Motivation for this article is two-fold. From the combinatorial side, we continue the recently started line of investigation of upper bounds on generalized additive energies of subsets A of the discrete cube

$$\{0,1,2,\ldots,n-1\}^d. \tag{1.1}$$

From the analytical side, we take omnipresent classical inequalities between the Gowers norms and the Lebesgue norms for functions $f\colon \mathbb{Z}^d \to \mathbb{C}$ and sharpen them when we additionally assume that f is supported in the cube (1.1).

The *additive energy*

$$E_2(A) := \big|\{(a_1,a_2,a_3,a_4) \in A^4 \,:\, a_1 - a_2 = a_3 - a_4\}\big| \tag{1.2}$$

of a finite set $A \subset \mathbb{Z}^d$ appears naturally in additive combinatorics; see the book by Tao and Vu [23, Section 2.3]. For sets $A \subseteq \{0,1\}^d \subset \mathbb{Z}^d$, Kane and Tao [16, Theorem 7] proved the optimal energy bound in terms of the set size,

$$E_2(A) \leqslant |A|^{\log_2 6}.$$

The exponent $\log_2 6$ is sharp since the equality holds when A is the whole binary cube $\{0,1\}^d$. Three generalizations of E_2 also appear naturally in applications and they were also systematically studied on their own by Schoen and Shkredov [20] and Shkredov [22]. De Dios Pont, Greenfeld, Ivanisvili, and Madrid [6] revisited two of them and called them the *higher energies*

$$\widetilde{E}_k(A) := \big|\{(a_1,a_2,\ldots,a_{2k-1},a_{2k}) \in A^{2k} \,:\, a_1 - a_2 = a_3 - a_4 = \cdots = a_{2k-1} - a_{2k}\}\big|$$

2020 *Mathematics Subject Classification.* Primary 05D05; Secondary 11B30.

Key words and phrases. additive energy, binary cube, Lebesgue norm, Shannon entropy, asymptotic formula.

and the *k-additive energies*

$$E_k(A) := \big|\{(a_1, a_2, \dots, a_{2k-1}, a_{2k}) \in A^{2k} \,:\, a_1 + \cdots + a_k = a_{k+1} + \cdots + a_{2k}\}\big|.$$

Here $k \geqslant 2$ is an integer and A is again a finite subset of $\mathbb{Z}^d$. Note that $\widetilde{E}_2 = E_2$ is the usual additive energy (1.2). The authors of [6] then also restricted their attention to the sets $A \subseteq \{0,1\}^d$. Following an inductive scheme similar to the one in [16, Section 2], they proved sharp inequalities for $\widetilde{E}_k$,

$$\widetilde{E}_k(A) \leqslant |A|^{\log_2(2^k+2)};$$

see [6, Theorem 1]. The energies E_k satisfy similarly looking sharp estimates

$$E_k(A) \leqslant |A|^{\log_2 \binom{2k}{k}},$$

but the proof was more technical: it was done for $k \leqslant 100$ in [6, Section 3], while the second author of the present paper established an ingredient needed for general k in [18]. A third possible quantity that generalizes (1.2) will be defined via the so-called Gowers norms in (1.3) below. It has already been studied by Shkredov [22], who also discussed numerous relations between all three kinds of generalized additive energies.

For a general complex function f defined on an additivelly written (discrete) abelian group and a positive integer k one defines the *Gowers uniformity norm* as

$$\|f\|_{\mathrm{U}^k} := \Big(\sum_{a,h_1,\dots,h_k} \prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k} \mathcal{C}^{\epsilon_1+\cdots+\epsilon_k} f(a + \epsilon_1 h_1 + \cdots + \epsilon_k h_k) \Big)^{1/2^k},$$

where $\mathcal{C}\colon z \mapsto \bar{z}$ denotes the operator of complex conjugation, so that $\mathcal{C}^2$ is the identity. This is, in fact, a norm for $k \geqslant 2$, defined on the set of functions for which the above expression is finite; see Section 2. These norms were introduced by Gowers [12, 13] in his quantitative proof of Szemerédi's theorem. If f is specialized to be the indicator function $\mathbb{1}_A$ of a finite set $A \subset \mathbb{Z}^d$, then the quantity

$$P_k(A) := \|\mathbb{1}_A\|_{\mathrm{U}^k}^{2^k} \tag{1.3}$$

rewritten as

$$P_k(A) = \big|\big\{(a, h_1, \dots, h_k) \in (\mathbb{Z}^d)^{k+1} \,:\, a + \epsilon_1 h_1 + \cdots + \epsilon_k h_k \in A \text{ for every } (\epsilon_1, \dots, \epsilon_k) \in \{0,1\}^k\big\}\big|$$

counts parallelotopes with vertices in A, with a lot of overcounting. Note that P_2 is again the ordinary additive energy E_2.

For a finitely supported function $f\colon \mathbb{Z}^d \to \mathbb{C}$ the inequality

$$\|f\|_{\mathrm{U}^k} \leqslant \|f\|_{\ell^p} \tag{1.4}$$

is well-known to hold for every exponent

$$p \leqslant \frac{2^k}{k+1}; \tag{1.5}$$

see [8]. Estimate (1.4), no matter how simple, has found versatile important applications, for instance in the study of generalized multiplicative sequences [10, Section 2] and automatic sequences [3, Proposition 2.2], or in geometric measure theory of large sets [7, Section 6]. The number $2^k/(k+1)$ is called the critical exponent in [8, Subsection 1.1] and it cannot be increased. Namely, it is easy to see that

$$\|\mathbb{1}_{\{0,1,\dots,n-1\}}\|_{\mathrm{U}^k} \text{ is comparable to } n^{(k+1)/2^k} \tag{1.6}$$

up to an unimportant multiplicative constant, while

$$\|\mathbb{1}_{\{0,1,\ldots,n-1\}}\|_{\ell^p} = n^{1/p}.$$

If we had (1.4) with some $p > 2^k/(k+1)$, we would arrive at a contradiction in the limit as $n \to \infty$.

One naturally wonders if (1.5) can still be improved if we additionally assume that f is supported in the binary cube. Also, one might regard P_k defined by (1.3) as yet another sequence of generalized energies of $A \subseteq \{0,1\}^d$ and attempt to control them by optimal powers of the size of A. Both of these problems are resolved by the first result of this paper.

Theorem 1. *Let $d \geqslant 0$ and $k \geqslant 2$ be integers. For every function $f\colon \mathbb{Z}^d \to \mathbb{C}$ supported in $\{0,1\}^d$ the inequality (1.4) holds for every exponent*

$$0 < p \leqslant \frac{2^k}{\log_2(2k+2)}.$$

In particular, by taking the largest possible p and $f = \mathbb{1}_A$, we get that

$$P_k(A) \leqslant |A|^{\log_2(2k+2)} \tag{1.7}$$

holds for every set $A \subseteq \{0,1\}^d$.

Both the stated range of p and inequality (1.7) are sharp since the equality holds for $A = \{0,1\}^d$. Note that $\mathbb{Z}^0$ is interpreted as a group consisting only of the neutral element for the addition.

Estimates for E_2 on larger discrete cubes (1.1) were studied by de Dios Pont, Greenfeld, Ivanisvili, and Madrid [6, Subsection 1.3] and Shao [21]; also see a more general setting by Hegyvári [14]. In this paper we also want to study nontrivial inequalities for $\|f\|_{\mathrm{U}^k}$ and $P_k(A)$, i.e., (1.4) and

$$P_k(A) \leqslant |A|^t, \tag{1.8}$$

on discrete cubes. The first step in these considerations is the following general proposition, which reduces higher-dimensional estimates to an optimization problem in finitely many variables. It will be formulated and proven in the spirit of the results from [6, Section 4].

Proposition 2. *Let $k, n \geqslant 2$ be integers and let $p, t > 0$ be real numbers such that $pt = 2^k$. The following are equivalent.*

1. *Inequality (1.4) holds for every function $f\colon \mathbb{Z} \to [0,\infty)$ supported in the set $\{0,1,2,\ldots,n-1\}$.*
2. *Inequality (1.4) holds for every integer $d \geqslant 0$ and every function $f\colon \mathbb{Z}^d \to \mathbb{C}$ supported in the discrete cube (1.1).*
3. *Inequality (1.8) holds for every integer $d \geqslant 0$ and every subset A of the discrete cube (1.1).*
4. *There exists a constant $C \in (0,\infty)$ such that $P_k(A) \leqslant C|A|^t$ holds for every integer $d \geqslant 0$ and every subset A of (1.1).*

All ingredients for the proof of Proposition 2 are borrowed from [6], even though the authors there rather discussed norms of multiple convolutions,

$$\|\underbrace{f * f * \cdots * f}_{k \text{ times}}\|_{\ell^2}^2$$

and the related k-additive energies E_k. Also the proofs in [6] crucially use the Fourier transform, while we avoid the need to look at the functions on the Fourier side at all. This is important here and in the later text, since the higher Gowers norms $\|f\|_{\mathrm{U}^k}$, $k \geqslant 3$, are not expressible simply as certain sizes of the Fourier transform of f.

Let $t_{k,n}$ denote the smallest number $t > 0$ such that (1.8) holds for every positive integer d and every subset A of (1.1). The smallest such number really exists by Proposition 2, since part (1) is structurally just an inequality for n non-negative numbers, which is a closed condition. By the same proposition we then also know that

$$p_{k,n} = 2^k / t_{k,n} \tag{1.9}$$

is the largest $p > 0$ such that (1.4) holds in any dimension d and for every function f supported in (1.1).

Theorem 1 can now be reformulated simply as

$$t_{k,2} = \log_2(2k+2)$$

for $k \geqslant 2$. It is not very likely that explicit expressions for $t_{k,n}$ can be computed when $n \geqslant 3$. Namely, similar computation to the one in [6, Subsection 4.3] shows that the exponent $t_{k,3}$ associated with the ternary cube $\{0,1,2\}^d$ is the smallest $t > 0$ such that

$$\begin{aligned} \max_{\substack{x,y,z\in[0,\infty)\\ x+y+z=1}} \big(& x^t + y^t + z^t + 2k x^{t/2} y^{t/2} + 2k y^{t/2} z^{t/2} \\ & + 2k x^{t/2} z^{t/2} + 2k(k-1) x^{t/4} y^{t/2} z^{t/4} \big) = 1. \end{aligned} \tag{1.10}$$

Already the number $t_{2,3}$ does not seem to be a “nice” explicit number. As remarked in [6, Subsection 4.3] its trivial bounds are

$$2.68014\ldots = \log_3 19 \leqslant t_{2,3} \leqslant 3,$$

where it is also shown that

$$t_{2,3} \geqslant 2\log_2 2.5664 = 2.71949\ldots.$$

Numerical computation in Mathematica [24] based on (1.10) and formula (1.9) give

$$t_{2,3} = 2.7207109973\ldots,$$
$$p_{2,3} = 1.4702039297\ldots.$$

Inverse Symbolic Calculator [2] does not recognize these two as any of the numbers appearing naturally elsewhere. Still, as a consequence of a general Theorem 4 below, we will have a quite precise asymptotics of $t_{k,3}$ for large k, namely

$$t_{k,3} = \frac{4}{3}\log_2 k + \frac{2}{3} + o^{k\to\infty}(1). \tag{1.11}$$

For the previous reason, the best one can hope for general $k \geqslant 2$ and $n \geqslant 3$ is to study the asymptotic behavior of $t_{k,n}$ as either $n \to \infty$ or $k \to \infty$. Let first fix k and study the asymptotics in n. As we have already mentioned, the numbers $t_{2,n}$ were studied before, since P_2 coincides with the ordinary additive energy (1.2). The bound

$$3 - \frac{\log_2 3 - 1}{\log_2 n} \leqslant t_{2,n} \leqslant 3$$

from [6, Proposition 7] was improved by Shao [21] to

$$3 - \big(1 + o^{n\to\infty}(1)\big)\frac{3\log_2 3 - 4}{2\log_2 n} \leqslant t_{2,n} \leqslant 3 - \frac{c}{\log_2 n} \tag{1.12}$$

for some constant $c > 0$. Shao also conjectured that the leftmost expression in (1.12) is the correct asymptotics of $t_{2,n}$ as $n \to \infty$. Here we generalize Shao's results to $t_{k,n}$ for $k \geqslant 3$.

Theorem 3. *There exists a constant $c \in (0, \infty)$ such that for every integer $k \geqslant 2$ we have*

$$k + 1 - \big(1 + o_k^{n\to\infty}(1)\big) \frac{(k+1)\log_2(k+1) - 2k}{2\log_2 n} \leqslant t_{k,n} \leqslant k + 1 - \frac{c}{\log_2 n}.$$

Note that the lower bound from Theorem 3 specializes to the one in (1.12) when $k = 2$ and one can again speculate that it is optimal. The constant c in the upper bound is the same one from (1.12), as we are, in fact, applying Shao's highly nontrivial bound as a black box. One could, in fact, repeat the proof from [21] to obtain constants c_k that strictly increase with k. However, we found it difficult to track down the actual growth and it is unlikely that it could match the growth of $(1/2)(k+1)\log_2(k+1) - k$ from the lower bound.

In this paper we also have the opportunity to fix $n \geqslant 2$ and study the asymptotics of $t_{k,n}$ as $k \to \infty$. Here is where we can give a quite definite result; we find it the most substantial contribution of this paper. Let

$$H_m := -\sum_{j=0}^{m} \frac{\binom{m}{j}}{2^m} \log_2 \frac{\binom{m}{j}}{2^m} \tag{1.13}$$

denote the entropy of the symmetric binomial distribution $B(m, 1/2)$; see Section 2 for the more general definition.

Theorem 4. *For every integer $n \geqslant 2$ we have*

$$t_{k,n} = \frac{(n-1)\log_2(2k) - \log_2(n-1)!}{H_{n-1}} + o_n^{k\to\infty}(1).$$

As a consequence, $t_{k,n}$ is asymptotically equal to

$$\frac{n-1}{H_{n-1}} \log_2 k$$

as $k \to \infty$. Several values of the leading coefficient $(n-1)/H_{n-1}$ are listed in Table 1. We also mention that it is asymptotically equal to $2n/\log_2 n$ as $n \to \infty$; this is a well-known fact, which also follows from sharper estimates (6.3) in Section 6. In particular, the constant $t_{k,3}$ relevant for the ternary cube is asymptotically equal to $(4/3)\log_2 k$ and Theorem 4 also gives the more precise formula (1.11). Somewhat surprisingly, the proof of Theorem 4 will require subtle estimates for the Shannon entropy of certain independent sums of discrete random variables, which will be discussed in Subsection 6.1. An interesting open problem along the same lines will be formulated too; see Conjecture 9.

2. Notation and preliminaries

Cardinality of a finite set S is simply written as $|S|$, just as we already did in the introduction. We write $\mathbb{N}$ for the set of positive integers, $\{1, 2, 3, \ldots\}$. Other number sets, namely $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$, have their usual meanings. Logarithms with different bases will be used in the text. By far the most commonly used will be 2. Logarithm with base e will be denoted ln.

n	$(n-1)/H_{n-1}$	numerical value
2	1	1
3	$4/3$	$1.3333333333\ldots$
4	$4/(4-\log_2 3)$	$1.6562889815\ldots$
5	$32/(21-3\log_2 3)$	$1.9698232317\ldots$
6	$16/(14-3\log_2 5)$	$2.2745961522\ldots$

TABLE 1. Several values of $(n-1)/H_{n-1}$.

Suppose that

$$F(m,a,b,\ldots)\in\mathbb{R} \quad\text{and}\quad G(m,a,b,\ldots)\in(0,\infty)$$

are quantities defined for sufficiently large positive integers m that possibly also depend on certain parameters $a,b,\ldots$. We write

$$F(m,a,b,\ldots)=O^{m\to\infty}_{a,b,\ldots}\big(G(m,a,b,\ldots)\big)$$

if

$$\limsup_{m\to\infty}\frac{|F(m,a,b,\ldots)|}{G(m,a,b,\ldots)}<\infty$$

for every possible $a,b,\ldots$, while

$$F(m,a,b,\ldots)=o^{m\to\infty}_{a,b,\ldots}\big(G(m,a,b,\ldots)\big)$$

means that

$$\lim_{m\to\infty}\frac{F(m,a,b,\ldots)}{G(m,a,b,\ldots)}=0$$

holds for every fixed choice of $a,b,\ldots$. We also simply write

$$O^{m\to\infty}_{a,b,\ldots}\big(G(m,a,b,\ldots)\big) \quad\text{and}\quad o^{m\to\infty}_{a,b,\ldots}\big(G(m,a,b,\ldots)\big)$$

in place of any particular expression $F(m,a,b,\ldots)$ with the above property. It is understood that these definitions are uniform over any other objects that do not appear in the subscript of the O or o notation. Next, we say that sequences $(F(m))_m$ and $(G(m))_m$ are asymptitocally equal if

$$\lim_{m\to\infty}\frac{F(m)}{G(m)}=1,$$

i.e.,

$$F(m)=(1+o^{m\to\infty}(1))\,G(m).$$

The ℓ^p *norm* of a function $f\colon\mathbb{Z}^d\to\mathbb{C}$ is defined as

$$\|f\|_{\ell^p}:=\Big(\sum_{x\in\mathbb{Z}^d}|f(x)|^p\Big)^{1/p}$$

for $p\in[1,\infty)$, while the ℓ^∞ *norm* is simply

$$\|f\|_{\ell^\infty}:=\sup_{x\in\mathbb{Z}^d}|f(x)|.$$

Convolution of functions $f,g\colon\mathbb{Z}^d\to\mathbb{C}$ is another complex function on $\mathbb{Z}^d$ defined as

$$(f*g)(x):=\sum_{y\in\mathbb{Z}^d}f(x-y)g(y)$$

for $x \in \mathbb{Z}^d$. *Young's inequality for convolution* now reads

$$\|f * g\|_{\ell^r} \leqslant \|f\|_{\ell^p} \|g\|_{\ell^q}, \tag{2.1}$$

which holds when $p, q, r \in [1, \infty]$ are such that $1/p + 1/q = 1 + 1/r$. We also define $\widetilde{f}$ to be the reflection of f, namely

$$\widetilde{f}(x) := f(-x)$$

for every $x \in \mathbb{Z}^d$, so that $\|\widetilde{f}\|_{\ell^p} = \|f\|_{\ell^p}$ for all $p \in [1, \infty]$.

The *Gowers "inner product"* and its properties will play a crucial role in the proofs given in Sections 3 and 4. It is defined to be

$$\begin{aligned} &\big\langle (f_{\epsilon_1,\ldots,\epsilon_k})_{(\epsilon_1,\ldots,\epsilon_k)\in\{0,1\}^k} \big\rangle_{\mathrm{U}^k} \\ &:= \sum_{a,h_1,\ldots,h_k} \prod_{(\epsilon_1,\ldots,\epsilon_k)\in\{0,1\}^k} \mathcal{C}^{\epsilon_1+\cdots+\epsilon_k} f_{\epsilon_1,\ldots,\epsilon_k}(a + \epsilon_1 h_1 + \cdots + \epsilon_k h_k), \end{aligned}$$

where $(f_{\epsilon_1,\ldots,\epsilon_k})$ is now a 2^k-tuple of complex functions on $\mathbb{Z}^d$ such that the above multiple series converges absolutely. It satisfies the so-called *Gowers–Cauchy–Schwarz inequality*:

$$\big|\big\langle (f_{\epsilon_1,\ldots,\epsilon_k})_{(\epsilon_1,\ldots,\epsilon_k)\in\{0,1\}^k} \big\rangle_{\mathrm{U}^k}\big| \leqslant \prod_{(\epsilon_1,\ldots,\epsilon_k)\in\{0,1\}^k} \|f_{\epsilon_1,\ldots,\epsilon_k}\|_{\mathrm{U}^k};$$

see [13, Lemma 3.8]. One if its consequences is that the Gowers norms satisfy the triangle inequality,

$$\|f_1 + f_2\|_{\mathrm{U}^k} \leqslant \|f_1\|_{\mathrm{U}^k} + \|f_2\|_{\mathrm{U}^k};$$

see [13, Lemma 3.9].

The notion of the Shannon entropy will be needed in Section 6. We only need to work with discrete probability distributions on the set $\{0, 1, \ldots, n-1\}$, namely

$$\begin{pmatrix} 0 & 1 & 2 & \cdots & n-1 \\ q_0 & q_1 & q_2 & \cdots & q_{n-1} \end{pmatrix} \tag{2.2}$$

for non-negative numbers $q_0, q_1, \ldots, q_{n-1}$ that sum to 1. If a random variable X has distribution (2.2), then its *Shannon entropy* (or simply just *entropy*) is defined to be the number

$$\mathrm{H}(X) := -\sum_{j=0}^{n-1} q_j \log_2 q_j. \tag{2.3}$$

Here we interpret $0 \log_2 0$ as 0. The number (2.3) will also sometimes be written as

$$\mathrm{H}(q_0, \ldots, q_{n-1}).$$

Recall that, in accordance with (1.13), the entropy of the symmetric binomial distribution with m trials, namely $B(m, 1/2)$, is also written simply as H_m.

3. Proof of Theorem 1

The proof outline is similar to the one used in [6, Sections 2&3]. First, we prove a generalization of an "elementary inequality" by Kane and Tao [16, Lemma 8], similar to but different from [6, Lemma 9], [6, Lemma 5], and [18, Theorem 1].

Fix an integer $k \geqslant 2$. For shortness we simply denote

$$t := \log_2(2k + 2)$$

throughout this section, always remembering that t depends on k.

Lemma 5. *For every* $x, y \in [0,\infty)$ *we have*

$$x^t + y^t + 2k x^{t/2} y^{t/2} \leqslant (x+y)^t.$$

Proof. The case $k = 2$ is precisely [16, Lemma 8], which has also been reproved several times in [6, 18]. Thus, we can assume that $k \geqslant 3$, which also guarantees $t \geqslant 3$.

The inequality is trivial for $x = y = 0$. In general, both sides of the inequality are homogeneous of order t in the variables x and y, so it is sufficient to prove it when we also normalize $x + y = 1$. Define the function $\varphi\colon [0,1] \to [0,\infty)$ as

$$\varphi(x) := x^t + (1-x)^t + 2k(x(1-x))^{t/2}.$$

Since $\varphi(1-x) = \varphi(x)$, we only need to show that

$$\varphi(x) \leqslant 1 \tag{3.1}$$

for every $x \in [0, 1/2]$. Let $x_0 \in \langle 0, 1/2\rangle$ be the unique solution $x = x_0$ to the equation

$$k(x(1-x))^{t/2-1} = 1,$$

which exists due to $k(1/4)^{t/2-1} > 1$. We prove (3.1) separately on two sub-intervals of $[0, 1/2]$.

Case 1: interval $[0, x_0]$. For every $0 \leqslant x \leqslant x_0$ we have, by the definition of x_0,

$$2k(x(1-x))^{t/2} \leqslant 2x(1-x),$$

which implies

$$\varphi(x) \leqslant x^2 + (1-x)^2 + 2x(1-x) = (x+1-x)^2 = 1$$

and so confirms (3.1).

Case 2: interval $\langle x_0, 1/2]$. Differentiating we get

$$\varphi'(x) = t x^{t-1} - t(1-x)^{t-1} + kt(x(1-x))^{t/2-1}(1-2x).$$

Also denote $\psi\colon [0,1/2] \to \mathbb{R}$,

$$\psi(x) := x^{t-1} - (1-x)^{t-1} + 1 - 2x.$$

Note that ψ is concave due to

$$\psi''(x) = (t-1)(t-2)\big(x^{t-3} - (1-x)^{t-3}\big) \leqslant 0.$$

Since, $\psi(0) = 0 = \psi(1/2)$, we also have that ψ is non-negative on $[0, 1/2]$. Now, for every $x \in \langle x_0, 1/2]$ we have

$$k(x(1-x))^{t/2-1} > 1,$$

which implies

$$\varphi'(x) > t\psi(x) \geqslant 0.$$

Consequently, φ is increasing on the interval $\langle x_0, 1/2]$, so

$$\varphi\Big(\frac{1}{2}\Big) = 2^{1-t} + 2k \cdot 2^{-t} = 1$$

guarantees that (3.1) holds on that interval too. ∎

Now we turn to the proof of Theorem 1. Desired inequality (1.4) in the endpoint case

$$p = \frac{2^k}{t}$$

can be rewritten as

$$\|f\|_{\mathrm{U}^k}^{2^k} \leqslant \big\||f|^p\big\|_{\ell^1}^t \tag{3.2}$$

and we can also safely assume that f is nonnegative. We proceed by the mathematical induction on d. The basis case $d = 0$ is trivial, since f is only defined at a single point. Take $d \geqslant 1$ and assume that (3.2) holds for functions on $\mathbb{Z}^{d-1}$ supported in $\{0,1\}^{d-1}$.

Now take an arbitrary function $f\colon \mathbb{Z}^d \to [0,\infty)$ supported in $\{0,1\}^d$ and define its "slices"

$$f_0, f_1\colon \mathbb{Z}^{d-1} \to [0,\infty)$$

by

$$f_0(x_1,\dots,x_{n-1}) := f(x_1,\dots,x_{n-1},0),$$
$$f_1(x_1,\dots,x_{n-1}) := f(x_1,\dots,x_{n-1},1)$$

for $(x_1,\dots,x_{n-1}) \in \mathbb{Z}^{d-1}$. The induction hypothesis applied to f_0 and f_1 gives

$$\|f_i\|_{\mathrm{U}^k}^{2^k} \leqslant \big\|f_i^p\big\|_{\ell^1}^t \tag{3.3}$$

for $i = 0,1$. Consider any $(k+1)$-tuple $(a, h_1, \dots, h_k)$ of vectors in $\mathbb{Z}^d$ such that

$$a + \epsilon_1 h_1 + \dots + \epsilon_k h_k \text{ falls into } \{0,1\}^d$$

for every choice of $(\epsilon_1,\dots,\epsilon_k) \in \{0,1\}^k$. Writing

$$a = (a', a''), \quad a' \in \mathbb{Z}^{d-1},\ a'' \in \mathbb{Z}$$

and

$$h_j = (h'_j, h''_j), \quad h'_j \in \mathbb{Z}^{d-1},\ h''_j \in \mathbb{Z}$$

for $j = 1,\dots,k$, we see that we do not have too many options for the values of $a'', h''_1, \dots, h''_k$. The possibilities are:

- $a'' = 0,\ h''_1 = \dots = h''_k = 0,$ or
- $a'' = 1,\ h''_1 = \dots = h''_k = 0,$ or
- $a'' = 0,\ h''_{j_0} = 1$ for some index j_0, $h''_j = 0$ for every index $j \neq j_0$, or
- $a'' = 1,\ h''_{j_0} = -1$ for some index j_0, $h''_j = 0$ for every index $j \neq j_0$.

Gathering all of the $2k+2$ possible cases we get

$$\begin{aligned}\|f\|_{\mathrm{U}^k}^{2^k} &= \|f_0\|_{\mathrm{U}^k}^{2^k} + \|f_1\|_{\mathrm{U}^k}^{2^k} + 2k\langle\underbrace{f_0,\dots,f_0}_{2^{k-1}},\underbrace{f_1,\dots,f_1}_{2^{k-1}}\rangle_{\mathrm{U}^k}\\ &\leqslant \|f_0\|_{\mathrm{U}^k}^{2^k} + \|f_1\|_{\mathrm{U}^k}^{2^k} + 2k\|f_0\|_{\mathrm{U}^k}^{2^{k-1}}\|f_1\|_{\mathrm{U}^k}^{2^{k-1}},\end{aligned}$$

where we have applied the Gowers–Cauchy–Schwarz inequality in the last line. By also applying (3.3) and Lemma 5 with

$$x = \big\|f_0^p\big\|_{\ell^1}, \quad y = \big\|f_1^p\big\|_{\ell^1}, \quad x + y = \big\|f^p\big\|_{\ell^1},$$

we obtain

$$\|f\|_{\mathrm{U}^k}^{2^k} \leqslant \big\|f_0^p\big\|_{\ell^1}^t + \big\|f_1^p\big\|_{\ell^1}^t + 2k\big\|f_0^p\big\|_{\ell^1}^{t/2}\big\|f_0^p\big\|_{\ell^1}^{t/2} \leqslant \big\|f^p\big\|_{\ell^1}^t,$$

which gives (3.2) and completes the induction step.

4. Proof of Proposition 2

Recall that p and t are now arbitrary positive numbers related by

$$pt = 2^k.$$

Proof of (1) $\Longrightarrow$ (2). The assumption (1) is that for every $g\colon \mathbb{Z}\to[0,\infty)$ supported in $\{0,1,\dots,n-1\}$ one has

$$\sum_{b,l_1,\dots,l_k\in\mathbb{Z}} \prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k} g(b+\epsilon_1 l_1+\cdots+\epsilon_k l_k) \leqslant \Big(\sum_{b=0}^{n-1} g(b)^p\Big)^{2^k/p}. \tag{4.1}$$

We prove the claim (2) by the induction on d. The basis case $d=0$ is trivial again, since all functions on $\mathbb{Z}^0$ are constants. Now take $d\in\mathbb{N}$ and a complex function f on $\mathbb{Z}^d$ supported in (1.1). For each $b\in\mathbb{Z}$ define

$$f_b\colon \mathbb{Z}^{d-1}\to\mathbb{C}, \quad f_b(a):=f(a,b)$$

for $a\in\mathbb{Z}^{d-1}$, and use the induction hypothesis applied to each of these functions to obtain

$$\|f_b\|_{\mathrm{U}^k} \leqslant \|f_b\|_{\ell^p}. \tag{4.2}$$

Note that f_b is identically zero unless $0\leqslant b\leqslant n-1$. By the definition of the Gowers norm we have

$$\begin{aligned}\|f\|_{\mathrm{U}^k}^{2^k} &= \sum_{\substack{a,h_1,\dots,h_k\in\mathbb{Z}^{d-1}\\ b,l_1,\dots,l_k\in\mathbb{Z}}} \prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k} \mathcal{C}^{\epsilon_1+\cdots+\epsilon_k} f_{b+\epsilon_1 l_1+\cdots+\epsilon_k l_k}(a+\epsilon_1 h_1+\cdots+\epsilon_k h_k)\\ &= \sum_{b,l_1,\dots,l_k\in\mathbb{Z}} \big\langle (f_{b+\epsilon_1 l_1+\cdots+\epsilon_k l_k})_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k}\big\rangle_{\mathrm{U}^k},\end{aligned}$$

so the Gowers–Cauchy–Schwarz inequality followed by (4.2) yields

$$\begin{aligned}\|f\|_{\mathrm{U}^k}^{2^k} &\leqslant \sum_{b,l_1,\dots,l_k\in\mathbb{Z}} \prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k} \|f_{b+\epsilon_1 l_1+\cdots+\epsilon_k l_k}\|_{\mathrm{U}^k}\\ &\leqslant \sum_{b,l_1,\dots,l_k\in\mathbb{Z}} \prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k} \|f_{b+\epsilon_1 l_1+\cdots+\epsilon_k l_k}\|_{\ell^p}.\end{aligned}$$

It remains to apply (4.1) with

$$g(b):=\|f_b\|_{\ell^p}$$

to conclude

$$\|f\|_{\mathrm{U}^k}^{2^k} \leqslant \Big(\sum_{b=0}^{n-1}\|f_b\|_{\ell^p}^p\Big)^{2^k/p} = \|f\|_{\ell^p}^{2^k},$$

which proves (2).

Proof of (2) $\Longrightarrow$ (3). This is obvious by taking $f=\mathbb{1}_A$, observing

$$\|\mathbb{1}_A\|_{\ell^p}^{2^k} = |A|^{2^k/p} = |A|^t,$$

and recalling the definition of P_k.

Proof of (3) $\Longrightarrow$ (4). This is obvious by taking $C=1$.

Proof of (4) $\Longrightarrow$ (1). The assumption (4) can be equivalently stated as: there exist a constant $D\in[1,\infty)$ such that

$$\|\mathbb{1}_A\|_{\mathrm{U}^k} \leqslant D|A|^{t/2^k} = D\|\mathbb{1}_A\|_{\ell^p} \tag{4.3}$$

for every $d \geqslant 0$ and every $A \subseteq \{0,1,\ldots,n-1\}^d$. The following idea is borrowed from [6, Proof of Lemma 19].

We will first prove that for every $d \in \mathbb{N}$ and every function $f\colon \mathbb{Z}^d \to [0,\infty)$ with support in (1.1) one has

$$\|f\|_{\mathrm{U}^k} \leqslant D(3 + d\log_2 n)\|f\|_{\ell^p}. \tag{4.4}$$

Denote $M := \|f\|_{\ell^\infty}$, $N = \lceil d\log_2 n\rceil$, and write

$$\frac{f(a)}{M} = \sum_{i=0}^{N} \frac{\beta_i(a)}{2^i} + \frac{f'(a)}{M}$$

for every $a \in \{0,1,\ldots,n-1\}^d$ and some $\beta_i(a) \in \{0,1\}$, $f'(a) \in [0, 2^{-N}M)$. For each index i define the set

$$A_i := \{a \,:\, \beta_i(a) = 1\}$$

and the function

$$f_i := \frac{M}{2^i}\mathbb{1}_{A_i} = \frac{M}{2^i}\beta_i.$$

An application of the estimate (4.3) to A_i yields

$$\|f_i\|_{\mathrm{U}^k} = \frac{M}{2^i}\|\mathbb{1}_{A_i}\|_{\mathrm{U}^k} \leqslant D\frac{M}{2^i}\|\mathbb{1}_{A_i}\|_{\ell^p} = D\|f_i\|_{\ell^p},$$

while f' is trivially controlled as

$$\|f'\|_{\mathrm{U}^k} \leqslant \frac{M}{2^N}\Big(\sum_{a,h_1,\ldots,h_k\in\{0,\ldots,n-1\}^d} 1\Big)^{1/2^k} \leqslant \frac{M}{2^N}(n^d)^{(k+1)/2^k} \leqslant M(n^d)^{(k+1)/2^k-1} \leqslant M.$$

Thus, the triangle inequality for the Gowers norm gives

$$\|f\|_{\mathrm{U}^k} = \Big\|\sum_{i=0}^{N} f_i + f'\Big\|_{\mathrm{U}^k} \leqslant \sum_{i=0}^{N}\|f_i\|_{\mathrm{U}^k} + \|f'\|_{\mathrm{U}^k} \leqslant D\Big(\sum_{i=0}^{N}\|f_i\|_{\ell^p} + \|f\|_{\ell^\infty}\Big).$$

Since $0 \leqslant f_i \leqslant f$, we clearly have $\|f_i\|_{\ell^p} \leqslant \|f\|_{\ell^p}$, which concludes

$$\|f\|_{\mathrm{U}^k} \leqslant D(N+2)\|f\|_{\ell^p}$$

and completes the proof of (4.4).

Finally, we use a tensoring trick to remove the constant. Take an arbitrary $g\colon \mathbb{Z} \to [0,\infty)$ supported in $\{0,1,\ldots,n-1\}$ and, for some $d \in \mathbb{N}$, define $f\colon \mathbb{Z}^d \to [0,\infty)$ to be the d-th tensor power of g, i.e.,

$$f(a_1,a_2,\ldots,a_d) := g(a_1)g(a_2)\cdots g(a_d).$$

Then

$$\|f\|_{\mathrm{U}^k} = \|g\|_{\mathrm{U}^k}^d \text{ and } \|f\|_{\ell^p} = \|g\|_{\ell^p}^d,$$

so taking the d-th roots of (4.4) gives

$$\|g\|_{\mathrm{U}^k} \leqslant \big(D(3 + d\log_2 n)\big)^{1/d}\|g\|_{\ell^p}.$$

Letting $d \to \infty$ we obtain

$$\|g\|_{\mathrm{U}^k} \leqslant \|g\|_{\ell^p}$$

and thus finalize the proof of (1).

5. Proof of Theorem 3

We separately prove the lower bound

$$\liminf_{n\to\infty}\big((t_{k,n}-k-1)\log_2 n\big) \geqslant -\frac{k+1}{2}\log_2(k+1)+k \tag{5.1}$$

and the upper bound

$$t_{k,n} \leqslant k+1-\frac{c}{\log_2 n}, \tag{5.2}$$

where $c>0$ is a constant such that (1.12) holds. Before we begin, note that

$$\|\mathbb{1}_{\{0,1,\dots,n-1\}}\|_{\mathrm{U}^k} \leqslant n^{1/p_{k,n}},$$

so (1.6) and (1.9) give trivial bounds

$$\frac{2^k}{k+1} \leqslant p_{k,n} \leqslant \frac{2^k}{k+1} + o_k^{n\to\infty}(1) \tag{5.3}$$

and

$$k+1-o_k^{n\to\infty}(1) \leqslant t_{k,n} \leqslant k+1. \tag{5.4}$$

5.1. **The lower bound.** The Gowers norm of a complex function f on the real line is defined as

$$\|f\|_{\mathrm{U}^k(\mathbb{R})} := \bigg(\int_{\mathbb{R}^{k+1}} \prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k} \mathcal{C}^{\epsilon_1+\cdots+\epsilon_k} f(x+\epsilon_1 h_1+\cdots+\epsilon_k h_k)\,\mathrm{d}x\,\mathrm{d}h_1\cdots\mathrm{d}h_k\bigg)^{1/2^k}.$$

By a result of Eisner and Tao [8, Theorem 1.12] the optimal constant C_k in the inequality

$$\|g\|_{\mathrm{U}^k(\mathbb{R})} \leqslant C_k \|g\|_{\mathrm{L}^{2^k/(k+1)}(\mathbb{R})}$$

on the real line equals

$$C_k = \frac{2^{k/2^k}}{(k+1)^{(k+1)/2^{k+1}}}. \tag{5.5}$$

The equality is attained for all Gaussian functions among other extremizers, namely, for $k\geqslant 2$, the Gaussians modulated by complex trigonometric polynomials of degree at most $k-1$. This gives us an idea to test inequality (1.4) against truncated discrete analogues of a Gaussian. This idea was already employed by Shao [21], who did not work with the Gowers norms, but rather motivated the choice by the fact that the Gaussians also extremize real line Young's convolution inequality.

For a parameter $M>1$ and an integer $n\geqslant 2$ define a function $f_{M,n}\colon \mathbb{Z}\to[0,\infty)$ by

$$f_{M,n}(m) := \begin{cases} \exp\Big(-4M^2\Big(\frac{m}{n}-\frac{1}{2}\Big)^2\Big) & \text{for } m\in\{0,1,2,\dots,n-1\},\\ 0 & \text{otherwise.}\end{cases}$$

We also introduce $g, g_M\colon\mathbb{R}\to[0,\infty)$ by the formulae

$$g(x) := e^{-x^2},$$
$$g_M(x) := \exp\Big(-4M^2\Big(x-\frac{1}{2}\Big)^2\Big) = g(2Mx-M).$$

Writing the $(k+1)$-dimensional integral as a limit of its Riemann sums we obtain

$$\lim_{n\to\infty}\frac{1}{n^{k+1}}\|f_{M,n}\|_{\mathrm{U}^k}^{2^k} = \|g_M\mathbb{1}_{[0,1]}\|_{\mathrm{U}^k(\mathbb{R})}^{2^k} = \frac{1}{(2M)^{k+1}}\|g\mathbb{1}_{[-M,M]}\|_{\mathrm{U}^k(\mathbb{R})}^{2^k}. \tag{5.6}$$

Also, (5.3) implies

$$\begin{aligned}\|f_{M,n}\|_{\ell^{p_{k,n}}}^{p_{k,n}} &= \sum_{m=0}^{n-1} \exp\Big(-4M^2 p_{k,n}\Big(\frac{m}{n}-\frac{1}{2}\Big)^2\Big) \\ &\leqslant \sum_{m=0}^{n-1} \exp\Big(-4M^2 \frac{2^k}{k+1}\Big(\frac{m}{n}-\frac{1}{2}\Big)^2\Big) = \|f_{M,n}\|_{\ell^{2^k/(k+1)}}^{2^k/(k+1)},\end{aligned}$$

so we similarly obtain

$$\begin{aligned}&\limsup_{n\to\infty} \frac{1}{n}\|f_{M,n}\|_{\ell^{p_{k,n}}}^{p_{k,n}} \leqslant \lim_{n\to\infty} \frac{1}{n}\|f_{M,n}\|_{\ell^{2^k/(k+1)}}^{2^k/(k+1)} \\ &= \|g_M \mathbb{1}_{[0,1]}\|_{\mathrm{L}^{2^k/(k+1)}(\mathbb{R})}^{2^k/(k+1)} = \frac{1}{2M}\|g\mathbb{1}_{[-M,M]}\|_{\mathrm{L}^{2^k/(k+1)}(\mathbb{R})}^{2^k/(k+1)}.\end{aligned}$$

Moreover, by (1.9) and (5.4),

$$\begin{aligned}&\limsup_{n\to\infty} \frac{1}{n^{t_{k,n}}}\|f_{M,n}\|_{\ell^{p_{k,n}}}^{2^k} \\ &= \limsup_{n\to\infty}\Big(\frac{1}{n}\|f_{M,n}\|_{\ell^{p_{k,n}}}^{p_{k,n}}\Big)^{t_{k,n}} = \Big(\limsup_{n\to\infty}\frac{1}{n}\|f_{M,n}\|_{\ell^{p_{k,n}}}^{p_{k,n}}\Big)^{\lim_{n\to\infty} t_{k,n}} \\ &= \Big(\limsup_{n\to\infty}\frac{1}{n}\|f_{M,n}\|_{\ell^{p_{k,n}}}^{p_{k,n}}\Big)^{k+1} \leqslant \frac{1}{(2M)^{k+1}}\|g\mathbb{1}_{[-M,M]}\|_{\mathrm{L}^{2^k/(k+1)}(\mathbb{R})}^{2^k}. \end{aligned} \tag{5.7}$$

Dividing (5.6) by (5.7) we conclude

$$\liminf_{n\to\infty} n^{t_{k,n}-k-1}\Big(\frac{\|f_{M,n}\|_{\mathrm{U}^k}}{\|f_{M,n}\|_{\ell^{p_{k,n}}}}\Big)^{2^k} \geqslant \Big(\frac{\|g\mathbb{1}_{[-M,M]}\|_{\mathrm{U}^k(\mathbb{R})}}{\|g\mathbb{1}_{[-M,M]}\|_{\mathrm{L}^{2^k/(k+1)}(\mathbb{R})}}\Big)^{2^k}.$$

The definition of $p_{k,n}$ guarantees

$$\|f_{M,n}\|_{\mathrm{U}^k} \leqslant \|f_{M,n}\|_{\ell^{p_{k,n}}},$$

so, taking logarithms, we obtain

$$\liminf_{n\to\infty}\big((t_{k,n}-k-1)\log_2 n\big) \geqslant 2^k \log_2 \frac{\|g\mathbb{1}_{[-M,M]}\|_{\mathrm{U}^k(\mathbb{R})}}{\|g\mathbb{1}_{[-M,M]}\|_{\mathrm{L}^{2^k/(k+1)}(\mathbb{R})}}.$$

Finally, we can take the limit as $M\to\infty$ of the right hand size and obtain

$$\liminf_{n\to\infty}\big((t_{k,n}-k-1)\log_2 n\big) \geqslant 2^k \log_2 \frac{\|g\|_{\mathrm{U}^k(\mathbb{R})}}{\|g\|_{\mathrm{L}^{2^k/(k+1)}(\mathbb{R})}} = 2^k \log_2 C_k,$$

which, by (5.5), is precisely (5.1).

5.2. **The upper bound.** We will show (5.2) by the induction on $k \geqslant 2$. The basis $k=2$ is precisely (1.12). The induction step follows from $t_{k+1,n} \leqslant t_{k,n}+1$, which is a clear consequence of the following implication: if

$$\|f\|_{\mathrm{U}^k}^{2^k} \leqslant \big\||f|^{2^k/t}\big\|_{\ell^1}^{t}, \tag{5.8}$$

holds for every function f supported on (1.1), then one also has

$$\|f\|_{\mathrm{U}^{k+1}}^{2^{k+1}} \leqslant \big\||f|^{2^{k+1}/(t+1)}\big\|_{\ell^1}^{t+1} \tag{5.9}$$

for every function f supported on the same cube. To verify the implication, we denote $g=|f|^{2^k/t}$. Applying (5.8) to the function $x\mapsto\overline{f(x+h)}f(x)$ for each fixed h and using Young's convolution inequality (2.1) gives

$$\begin{aligned}\|f\|_{\mathrm{U}^{k+1}}^{2^{k+1}} &= \sum_{h\in\mathbb{Z}^d}\big\|\overline{f(\cdot+h)}f(\cdot)\big\|_{\mathrm{U}^k}^{2^k} \leqslant \sum_{h\in\mathbb{Z}^d}\big\||f(\cdot+h)|^{2^k/t}|f(\cdot)|^{2^k/t}\big\|_{\ell^1}^t\\ &= \|g*\widetilde{g}\|_{\ell^t}^t \leqslant \|g\|_{\ell^{2t/(t+1)}}^t\|\widetilde{g}\|_{\ell^{2t/(t+1)}}^t = \|g\|_{\ell^{2t/(t+1)}}^{2t} = \big\||f|^{2^{k+1}/(t+1)}\big\|_{\ell^1}^{t+1}.\end{aligned}$$

Thus, (5.8) implies (5.9).

6. Proof of Theorem 4

Before we begin with the proof of Theorem 4, we need to formulate and prove several auxiliary inequalities for the Shannon entropy of special distributions, some of which we could not find in the literature. Afterwards, we will separately prove the lower bound

$$\liminf_{k\to\infty}\Big(t_{k,n}-\frac{(n-1)\log_2(2k)-\log_2(n-1)!}{H_{n-1}}\Big)\geqslant 0 \tag{6.1}$$

and the upper bound

$$\limsup_{k\to\infty}\Big(t_{k,n}-\frac{(n-1)\log_2(2k)-\log_2(n-1)!}{H_{n-1}}\Big)\leqslant 0. \tag{6.2}$$

6.1. **Entropy estimates.** Let $X_1,X_2,X_3,\ldots$ be independent symmetric Bernoulli random variables, namely

$$X_i\sim\begin{pmatrix}0 & 1\\ 1/2 & 1/2\end{pmatrix},$$

so that $X_1+X_2+\cdots+X_m\sim B(m,1/2)$.

Lemma 6. *For every $m\in\mathbb{N}$ and arbitrary nonzero $h_1,\ldots,h_m\in\mathbb{Z}$ we have*

$$\mathrm{H}(h_1X_1+\cdots+h_mX_m)>\frac{1}{2}\log_2\frac{\pi m}{2}.$$

Proof. Fix a positive integer m and non-zero integers $h_1,\ldots,h_m$. Denote

$$X:=h_1X_1+\cdots+h_mX_m.$$

By a classical result of Erdős on the Littlewood–Offord problem [9, Theorem 1] we know that

$$\max_{i\in\mathbb{Z}}\mathbb{P}(X=i)\leqslant\binom{m}{\lfloor m/2\rfloor}2^{-m}.$$

The inequality

$$\binom{m}{\lfloor m/2\rfloor}2^{-m}<\sqrt{\frac{2}{\pi m}}$$

is well-known and it can be found in [19, p. 466]. Combining these two estimates we obtain

$$\mathrm{H}(X)=\sum_{i\in\mathbb{Z}}\mathbb{P}(X=i)\log_2\frac{1}{\mathbb{P}(X=i)}>\sum_{i\in\mathbb{Z}}\mathbb{P}(X=i)\log_2\sqrt{\frac{\pi m}{2}}=\frac{1}{2}\log_2\frac{\pi m}{2},$$

as desired. ∎

We will also need more precise estimates in the particular case of the entropy

$$H_m = \mathrm{H}(X_1 + \cdots + X_m)$$

of the symmetric binomial distribution $B(m, 1/2)$, which is the same number that was introduced in (1.13). The asymptotic series for the numbers H_m as $m \to \infty$ has been derived by several authors [15, 17, 11]. However, we need sharp exact inequalities for these numbers. They are available from the work of Adell, Lekuona, and Yu [1], who strengthened the previous bounds by Chang and Weldon [4].

Lemma 7 ([1])**.** *We have*

$$\frac{1}{2}\log_2 \frac{e\pi m}{2} - \frac{1}{4m} < H_m < \frac{1}{2}\log_2 \frac{e\pi m}{2} + \frac{1}{10m} \tag{6.3}$$

for every $m \in \mathbb{N}$.

Proof. The authors of [1] use the natural logarithm in the definition of the entropy (i.e., they measure it in nats), whereas we chose the base-2-logarithm in (2.3) (i.e., the entropy is measured in bits). Apart from that modification, [1, Formula (7)], which is a consequence of [1, Corollary 1], specialized to the symmetric binomial distribution reads

$$\frac{C_1}{m} + \frac{C_2}{m^2} + \frac{C_3}{m^3} < H_m - \frac{1}{2}\log_2 \frac{e\pi m}{2} < \frac{C_4}{m},$$

where

$$\begin{aligned} C_1 &= -0.24606\ldots, \\ C_2 &= 0.17527\ldots, \\ C_3 &= -0.00400\ldots, \\ C_4 &= 0.08202\ldots. \end{aligned}$$

It remains to disregard the terms C_2/m^2 and C_3/m^3 on the left hand side, since their sum is always positive. ☐

Finally, we can formulate and prove an inequality for the Shannon entropy that will play a crucial role in the proof of Theorem 4.

Lemma 8. *For every* $n \geqslant 2$, $1 \leqslant l \leqslant n-1$, *and* $h_1, \ldots, h_l \in \mathbb{Z} \setminus \{0\}$ *such that* $|h_1| + \cdots + |h_l| \leqslant n-1$ *we have*

$$\frac{\mathrm{H}(h_1X_1 + \cdots + h_lX_l)}{l} \geqslant \frac{H_{n-1}}{n-1}, \tag{6.4}$$

with equality attained only when $l = n-1$ *and* $|h_1| = \cdots = |h_{n-1}| = 1$, *in which case* $h_1X_1 + \cdots + h_{n-1}X_{n-1} \sim B(n-1, 1/2)$.

Proof. Changing some signs among $h_1, \ldots, h_l$ only translates the distribution of $h_1X_1 + \cdots + h_lX_l$ along the integers and preserves the entropy, because

$$-h_iX_i = h_i(1 - X_i) - h_i \sim h_iX_i - h_i.$$

Thus, we can assume that all numbers h_i are positive integers. After this reduction, the case $l = n-1$ leads to the unique possibility $h_1 = \cdots = h_{n-1} = 1$ and to the distribution $B(n-1, 1/2)$. Therefore, it remains to prove the strict inequality in (6.4) for $1 \leqslant l \leqslant n-2$. This also ensures that $n \geqslant 3$.

In the following proof we will first assume that $n > 100$ and then handle the cases $n \leqslant 100$ using a computer. Moreover, for large n we will distinguish two further subcases, depending on the magnitude of l with respect to n.

Case 1: $n > 100$ and $1 \leqslant l \leqslant 3(n-1)/4$. From Lemmas 6 and 7 respectively, we obtain

$$\begin{aligned}&2(n-1)\mathrm{H}(h_1X_1+\cdots+h_lX_l)-2lH_{n-1}\\&>(n-1)\log_2\frac{\pi l}{2}-2lH_{n-1}\\&>(n-1)\log_2\frac{\pi l}{2}-l\log_2\frac{e\pi(n-1)}{2}-\frac{l}{5(n-1)}.\end{aligned}$$

The function

$$\varphi(x):=(n-1)\log_2\frac{\pi x}{2}-x\log_2\frac{e\pi(n-1)}{2}-\frac{x}{5(n-1)}$$

has the second derivative

$$\varphi''(x)=-\frac{n-1}{x^2\ln 2}<0.$$

Thus, it is concave on $[1,3(n-1)/4]$, so it remains to verify its positivity at the endpoints:

$$\varphi(1)>0 \quad \text{for } n>100, \tag{6.5}$$

$$\varphi\Big(\frac{3(n-1)}{4}\Big)>0 \quad \text{for } n>100. \tag{6.6}$$

From

$$\log_2 s<\frac{s}{40} \quad \text{for } s\geqslant 50e\pi$$

we conclude

$$\begin{aligned}\varphi(1)&=(n-1)\log_2\frac{\pi}{2}-\log_2\frac{e\pi(n-1)}{2}-\frac{1}{5(n-1)}\\&>\frac{n-1}{2}-\frac{n-1}{5}-1>0,\end{aligned}$$

which proves (6.5). Also, to prove (6.6) we only need to observe

$$\begin{aligned}\varphi\Big(\frac{3(n-1)}{4}\Big)&=\frac{n-1}{4}\log_2(n-1)+(n-1)\Big(\log_2\frac{3\pi}{8}-\frac{3}{4}\log_2\frac{e\pi}{2}\Big)-\frac{3}{20}\\&>\frac{n-1}{4}\cdot 6-(n-1)-\frac{3}{20}>0.\end{aligned}$$

Case 2: $n>100$ and $3(n-1)/4<l\leqslant n-2$. Denote

$$u:=\big|\big\{i\in\{1,\ldots,l\}:h_i=1\big\}\big|,\quad v:=\big|\big\{i\in\{1,\ldots,l\}:h_i\geqslant 2\big\}\big|.$$

Clearly,

$$u+v=l,\quad u+2v\leqslant\sum_{i=1}^{l}h_i\leqslant n-1,$$

which gives

$$u=2(u+v)-(u+2v)\geqslant 2l-n+1.$$

Since the numbers h_i can be freely permuted without changing the distribution of $h_1X_1+\cdots+h_lX_l$, without loss of generality we assume that

$$h_1=\cdots=h_u=1.$$

For independent integer-valued random variables X and Y we have

$$\mathrm{H}(X+Y)\geqslant\mathrm{H}(X). \tag{6.7}$$

This follows from basic properties of the entropy and the conditional entropy [5, Chapter 2]. Alternatively, one can simply use Jensen's inequality and concavity of the function $x \mapsto -x \log_2 x$ on $[0,1]$: if $\mathbb{P}(X=i)=q_i$ and $\mathbb{P}(Y=i)=r_i$, then

$$-\sum_i \Big(\sum_j q_{i-j} r_j\Big) \log_2 \Big(\sum_j q_{i-j} r_j\Big) \geqslant \sum_i \sum_j r_j \big(-q_{i-j} \log_2 q_{i-j}\big) = -\sum_i q_i \log_2 q_i.$$

Applying (6.7) to

$$X := X_1 + \cdots + X_{2l-n+1}, \quad Y := h_{2l-n+2} X_{2l-n+2} + \cdots + h_l X_l$$

gives

$$\mathrm{H}(h_1 X_1 + \cdots + h_l X_l) \geqslant H_{2l-n+1}.$$

By Lemma 7 we now have

$$\begin{aligned} &2(n-1)\mathrm{H}(h_1 X_1 + \cdots + h_l X_l) - 2l H_{n-1} \\ &\geqslant 2(n-1) H_{2l-n+1} - 2l H_{n-1} \\ &> (n-1)\log_2 \frac{e\pi(2l-n+1)}{2} - \frac{n-1}{2(2l-n+1)} - l \log_2 \frac{e\pi(n-1)}{2} - \frac{l}{5(n-1)}. \end{aligned}$$

The function

$$\psi(x) := (n-1)\log_2 \frac{e\pi(2x-n+1)}{2} - \frac{n-1}{2(2x-n+1)} - x \log_2 \frac{e\pi(n-1)}{2} - \frac{x}{5(n-1)}$$

satisfies

$$\psi''(x) = -\frac{4(n-1)}{(2x-n+1)^2 \ln 2} - \frac{4(n-1)}{(2x-n+1)^3} < 0,$$

so it is clearly concave on $[3(n-1)/4, n-2]$. Therefore, it again remains to verify that it is positive at the endpoints:

$$\psi\Big(\frac{3(n-1)}{4}\Big) > 0 \quad \text{for } n > 100, \tag{6.8}$$

$$\psi(n-2) > 0 \quad \text{for } n > 100. \tag{6.9}$$

The proof of (6.8) is easy:

$$\begin{aligned} \psi\Big(\frac{3(n-1)}{4}\Big) &= \frac{n-1}{4}\log_2(n-1) + (n-1)\Big(\frac{1}{4}\log_2(e\pi) - \frac{5}{4}\Big) - \frac{23}{20} \\ &> \frac{n-1}{4}\cdot 6 - \frac{1}{2}(n-1) - \frac{23}{20} > 0, \end{aligned}$$

while

$$\begin{aligned} \psi(n-2) &= \log_2 \frac{e\pi(n-3)}{2} - (n-2)\log_2 \frac{n-1}{n-3} - \frac{n-1}{2(n-3)} - \frac{n-2}{5(n-1)} \\ &> 8 - 3 - 1 > 0 \end{aligned}$$

proves (6.9).

Case 3: $12 \leqslant n \leqslant 100$. From the previous cases we see that it is sufficient to make sure that for every $1 \leqslant l \leqslant n-2$ we have

$$(n-1)\log_2 \frac{\pi l}{2} - 2l H_{n-1} > 0$$

or

$$2(n-1) H_{2l-n+1} - 2l H_{n-1} > 0.$$

(In the second inequality it is understood that we must also have $l > (n-1)/2$.) Exact symbolic computation in Mathematica [24] verifies that, for every such l, at least one of the two inequalities holds, as soon as $12 \leqslant n \leqslant 100$.

Case 4: $3 \leqslant n \leqslant 11$. For every such n there are only finitely many cases of h_i to check the original claim of the lemma. Thus the verification can again be easily done using the symbolic computation in Mathematica [24]. Note that this type of checking would be too extensive for all $n \leqslant 100$. □

Let us also mention a conjecture on when

$$\mathrm{H}(h_1X_1 + \cdots + h_mX_m)$$

is minimized, given that the number of terms m is fixed. We arrived at it while trying to find a more streamlined proof of Lemma 8.

Conjecture 9. *For every $m \in \mathbb{N}$ and arbitrary $h_1, \ldots, h_m \in \mathbb{Z} \setminus \{0\}$ we have*

$$\mathrm{H}(h_1X_1 + \cdots + h_mX_m) \geqslant H_m,$$

with equality attained only when $|h_1| = \cdots = |h_m|$, in which case $h_1X_1 + \cdots + h_mX_m$ has the symmetric binomial distribution $B(m, 1/2)$.

We do not see immediate applications of Conjecture 9, provided that it turns out to be true, but we believe that it could be interesting on its own. Comparing it with the easy Lemma 6 and its proof, one can think of Conjecture 9 as an entropic variant of the Littlewood–Offord problem. A possible reason why it has not been studied before is that it asks to prove a sharp inequality for every fixed $m \in \mathbb{N}$, while it is trivial asymptotically as $m \to \infty$.

6.2. **The lower bound.** Recall that, by Proposition 2, the number $t_{k,n}$ is the smallest $t > 0$ such that

$$\sum_{a,h_1,\ldots,h_k\in\mathbb{Z}} \prod_{(\epsilon_1,\ldots,\epsilon_k)\in\{0,1\}^k} f(a + \epsilon_1h_1 + \cdots + \epsilon_kh_k) \leqslant \Big(\sum_{j=0}^{n-1} f(j)^{2^k/t}\Big)^t \tag{6.10}$$

holds for every function $f\colon \mathbb{Z} \to [0,\infty)$ supported in $\{0, 1, \ldots, n-1\}$. On the left hand side of inequality (6.10) we only observe the mutually equal terms obtained by taking:

- $a \in \{0, 1, \ldots, n-1\}$ arbitrary,
- precisely a of the numbers $h_1, \ldots, h_k$ to be equal -1,
- precisely $n-1-a$ of the numbers $h_1, \ldots, h_k$ to be equal 1, and
- precisely $k-n+1$ of the numbers $h_1, \ldots, h_k$ to be equal 0.

There are

$$\binom{k}{n-1} \sum_{a=0}^{n-1} \binom{n-1}{a} = \binom{k}{n-1} 2^{n-1}$$

such choices altogether, and they all contribute to the sum in (6.10) with the same term

$$\prod_{j=0}^{n-1} f(j)^{\binom{n-1}{j}2^{k-n+1}}.$$

Thus, by using inequality (6.10) with $t = t_{k,n}$, we obtain

$$\binom{k}{n-1} 2^{n-1} \prod_{j=0}^{n-1} f(j)^{\binom{n-1}{j}2^{k-n+1}} \leqslant \Big(\sum_{j=0}^{n-1} f(j)^{2^k/t_{k,n}}\Big)^{t_{k,n}}.$$

Now take a particular function f defined as

$$f(j) := \Big(\frac{\binom{n-1}{j}}{2^{n-1}}\Big)^{t_{k,n}/2^k}$$

for $0 \leqslant j \leqslant n-1$ and $f(j) := 0$ otherwise, so that the last inequality gives us

$$\binom{k}{n-1} 2^{n-1} \prod_{j=0}^{n-1} \Big(\frac{\binom{n-1}{j}}{2^{n-1}}\Big)^{\binom{n-1}{j} t_{k,n}/2^{n-1}} \leqslant \Big(\sum_{j=0}^{n-1} \frac{\binom{n-1}{j}}{2^{n-1}}\Big)^{t_{k,n}} = 1.$$

Taking logarithms,

$$\sum_{j=0}^{n-2} \log_2(k-j) - \log_2(n-1)! + (n-1) + \underbrace{\sum_{j=0}^{n-1} \frac{\binom{n-1}{j} t_{k,n}}{2^{n-1}} \log_2 \frac{\binom{n-1}{j}}{2^{n-1}}}_{-H_{n-1}t_{k,n}} \leqslant 0,$$

i.e.,

$$\liminf_{k\to\infty} \big(H_{n-1}t_{k,n} - (n-1)\log_2(2k) + \log_2(n-1)!\big) \geqslant \lim_{k\to\infty} \sum_{j=0}^{n-2} \log_2(1 - j/k) = 0,$$

which proves (6.1).

6.3. **The upper bound.** Note that the term on the left hand side of (6.10) corresponding to at least n nonzero numbers $h_1, \ldots, h_k$ must be identically 0. At the other extreme are the terms when $h_1 = \cdots = h_k = 0$. All remaining terms can be grouped by choosing $1 \leqslant l \leqslant n-1$ nonzero numbers among $h_1, \ldots, h_k$ in $\binom{k}{l}$ ways. Therefore, for a fixed integer $n \geqslant 2$ and every $l \in \{1, \ldots, n-1\}$ we define

$$T_{n,l} := \big\{(a, h_1, \ldots, h_l) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})^l \ : 0 \leqslant a + \epsilon_1 h_1 + \cdots + \epsilon_l h_l \leqslant n-1 \text{ for every } (\epsilon_1, \ldots, \epsilon_l) \in \{0,1\}^l\big\},$$

so that inequality (6.10) can now be rewritten as

$$\sum_{j=0}^{n-1} f(j)^{2^k} + \sum_{l=1}^{n-1} \sum_{(a,h_1,\ldots,h_l)\in T_{n,l}} \binom{k}{l} \prod_{(\epsilon_1,\ldots,\epsilon_l)\in\{0,1\}^l} f(a + \epsilon_1 h_1 + \cdots + \epsilon_l h_l)^{2^{k-l}} \leqslant \Big(\sum_{j=0}^{n-1} f(j)^{2^k/t}\Big)^t.$$

By substituting $g(j) = f(j)^{2^k/t}$, the estimate (6.10) is further equivalent to

$$\sum_{j=0}^{n-1} g(j)^t + \sum_{l=1}^{n-1} \sum_{(a,h_1,\ldots,h_l)\in T_{n,l}} \binom{k}{l} \prod_{(\epsilon_1,\ldots,\epsilon_l)\in\{0,1\}^l} g(a + \epsilon_1 h_1 + \cdots + \epsilon_l h_l)^{t/2^l} \leqslant 1 \qquad (6.11)$$

for every $g\colon \{0, 1, \ldots, n-1\} \to [0, \infty)$ such that $\sum_{j=0}^{n-1} g(j) = 1$. It is understood that only such functions g are considered in the rest of this subsection. We will prove (6.11) for any given $0 < \delta < 1$, for

$$t = \frac{(n-1)\log_2(2k) - \log_2(n-1)!}{H_{n-1}} + \delta, \qquad (6.12)$$

and for all positive integers k that are sufficiently large depending on n and δ. This will establish (6.2) by showing that the upper limit in question is at most δ for every $\delta > 0$.

Let us first make some preliminary observations. Every summand, other than $g(j)^t$, on the left hand side of (6.11) is of the form

$$\binom{k}{l}\big(g(0)^{q_0}g(1)^{q_1}\cdots g(n-1)^{q_{n-1}}\big)^t \tag{6.13}$$

where $(q_0, q_1, \ldots, q_{n-1})$ is some discrete probability distribution on the set $\{0, 1, \ldots, n-1\}$, which is also the distribution of some random variable $h_1X_1+\cdots+h_lX_l$ for $(a, h_1, \ldots, h_l) \in T_{n,l}$. We always interpret $0^0 = 1$. For every g as before the inequality between the weighted arithmetic and geometric means gives

$$1 = \sum_{j=0}^{n-1} g(j) \geqslant \sum_{\substack{0\leqslant j\leqslant n-1\\ q_j\neq 0}} q_j \frac{g(j)}{q_j} \geqslant \prod_{\substack{0\leqslant j\leqslant n-1\\ q_j\neq 0}} \Big(\frac{g(j)}{q_j}\Big)^{q_j} = \frac{\prod_{j=0}^{n-1} g(j)^{q_j}}{\prod_{j=0}^{n-1} q_j^{q_j}}.$$

That way we have obtained

$$\prod_{j=0}^{n-1} g(j)^{q_j} \leqslant 2^{-\mathrm{H}(q_0,\ldots,q_{n-1})}, \tag{6.14}$$

where we recall that $\mathrm{H}(q_0, \ldots, q_{n-1})$ is defined by the formula (2.3). Denote

$$\theta = 2^{-n2^n H_{n-1}/(n-1)}.$$

The proof of (6.11) is split into two cases. Once again, we always assume that g attains non-negative values that sum to 1.

Case 1: for every $0 \leqslant j \leqslant n-1$ *we have* $g(j) \leqslant 1-\theta$. Each summand (6.13) from the left hand side of (6.11) corresponding to some $(a, h_1, \ldots, h_l) \in T_{n,l}$ is, by inequality (6.14), at most

$$\binom{k}{l} 2^{-\mathrm{H}(h_1X_1+\cdots+h_lX_l)t},$$

which is, by the choice (6.12) for t, less than or equal to

$$\begin{cases} 2^{-n+1}\cdot 2^{-H_{n-1}\delta} & \text{for } l = n-1, \qquad (6.15)\\ O_n^{k\to\infty}\big(k^{l-(n-1)\mathrm{H}(h_1X_1+\cdots+h_lX_l)/H_{n-1}}\big) & \text{for } 1\leqslant l\leqslant n-2. \qquad (6.16)\end{cases}$$

In (6.15) we used the fact that the only possibility to have $l = n-1$ is $h_1X_1 + \cdots + h_{n-1}X_{n-1} \sim B(n-1, 1/2)$. Since $T_{n,n-1}$ has precisely 2^{n-1} elements described in the previous subsection and each corresponding term is bounded by (6.15), we conclude

$$\sum_{(a,h_1,\ldots,h_{n-1})\in T_{n,n-1}} \binom{k}{n-1} \prod_{(\epsilon_1,\ldots,\epsilon_{n-1})\in\{0,1\}^{n-1}} g(a+\epsilon_1h_1+\cdots+\epsilon_{n-1}h_{n-1})^{t/2^{n-1}} \leqslant 2^{-H_{n-1}\delta}.$$

Next, each term corresponding to $T_{n,l}$, $1 \leqslant l \leqslant n-2$, is bounded by (6.16) and the exponent of k is strictly negative by Lemma 8. Thus,

$$\sum_{l=1}^{n-2} \sum_{(a,h_1,\ldots,h_l)\in T_{n,l}} \binom{k}{l} \prod_{(\epsilon_1,\ldots,\epsilon_l)\in\{0,1\}^l} g(a+\epsilon_1h_1+\cdots+\epsilon_lh_l)^{t/2^l} = o_n^{k\to\infty}(1).$$

Finally, under the standing assumptions on g and by formula (6.12),

$$\sum_{j=0}^{n-1} g(j)^t \leqslant n(1-\theta)^t = o_n^{k\to\infty}(1).$$

Altogether, the left hand side of (6.11) is bounded by

$$\underbrace{2^{-H_{n-1}\delta}}_{<1} + o_n^{k\to\infty}(1)$$

for all functions g considered in this case. This is certainly less than 1 for sufficiently large k, just as desired.

Case 2: for some $0 \leqslant j_0 \leqslant n-1$ *we have* $g(j_0) > 1-\theta$. As a consequence of $\sum_j g(j) = 1$ we also have $g(j) < \theta$ for each $j \neq j_0$. Denote

$$N := \sum_{l=1}^{n-1} |T_{n,l}|.$$

Take k sufficiently large that $t > 2^n$; a further largeness requirement on k will be imposed later. The terms $g(j)^t$ in (6.11) are here simply controlled as

$$g(j)^t \leqslant g(j)^2$$

for $j = 0, 1, \ldots, n-1$. Each of the remaining N terms in (6.11) is at most

$$k^{n-1} g(j_1)^{t/2^n} g(j_2)^{t/2^n} \tag{6.17}$$

for some distinct indices j_1 and j_2. Since at least one of these indices is different from j_0, we can bound (6.17) from the above by

$$\begin{aligned} g(j_1)g(j_2)k^{n-1}\theta^{t/2^n-1} &= g(j_1)g(j_2)k^{n-1}O_n^{k\to\infty}(k^{-n}) \\ &= g(j_1)g(j_2)O_n^{k\to\infty}(k^{-1}) \leqslant \frac{2}{N} g(j_1)g(j_2), \end{aligned}$$

where the last inequality holds as soon as k is large enough depending on n. Altogether, the left hand side of (6.11) is then less than or equal to

$$\sum_{j=0}^{n-1} g(j)^2 + N\cdot\frac{2}{N}\max_{j_1\neq j_2} g(j_1)g(j_2) \leqslant \Big(\sum_{j=0}^{n-1} g(j)\Big)^2 = 1.$$

Both cases are now complete and they finalize the proof of inequality (6.11), and thus also of the upper bound (6.2).

Acknowledgment

This work was supported in part by the Croatian Science Foundation under the project HRZZ-IP-2022-10-5116 (FANAP).

References

[1] José A. Adell, Alberto Lekuona, and Yaming Yu. Sharp bounds on the entropy of the Poisson law and related quantities. *IEEE Trans. Inform. Theory*, 56(5):2299–2306, 2010. doi:10.1109/TIT.2010.2044057.

[2] Peter Benjamin Borwein, Jonathan Michael Borwein, and Simon Plouffe. Inverse symbolic calculator, 1995. URL: https://wayback.cecm.sfu.ca/projects/ISC/ISCmain.html.

[3] Jakub Byszewski, Jakub Konieczny, and Clemens Müllner. Gowers norms for automatic sequences. *Discrete Anal.*, pages Paper No. 4, 62, 2023.

[4] Shih Chun Chang and Edward J. Weldon, Jr. Coding for T-user multiple-access channels. *IEEE Trans. Inform. Theory*, 25(6):684–691, 1979. doi:10.1109/TIT.1979.1056109.

[5] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006.

[6] Jaume de Dios Pont, Rachel Greenfeld, Paata Ivanisvili, and José Madrid. Additive energies on discrete cubes. *Discrete Anal.*, pages Paper No. 13, 16, 2023. doi:10.19086/da.84737.
[7] Polona Durcik and Vjekoslav Kovač. A Szemerédi-type theorem for subsets of the unit cube. *Anal. PDE*, 15(2):507–549, 2022. doi:10.2140/apde.2022.15.507.
[8] Tanja Eisner and Terence Tao. Large values of the Gowers-Host-Kra seminorms. *J. Anal. Math.*, 117:133–186, 2012. doi:10.1007/s11854-012-0018-2.
[9] Paul Erdős. On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc.*, 51:898–902, 1945. doi:10.1090/S0002-9904-1945-08454-7.
[10] Aihua Fan and Jakub Konieczny. On uniformity of q-multiplicative sequences. *Bull. Lond. Math. Soc.*, 51(3):466–488, 2019. doi:10.1112/blms.12245.
[11] Philippe Flajolet. Singularity analysis and asymptotics of Bernoulli sums. *Theoret. Comput. Sci.*, 215(1-2):371–381, 1999. doi:10.1016/S0304-3975(98)00220-5.
[12] William Timothy Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998. doi:10.1007/s000390050065.
[13] William Timothy Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001. doi:10.1007/s00039-001-0332-9.
[14] Norbert Hegyvári. Some remarks on the distribution of additive energy. *Results Math.*, 78(6):Paper No. 233, 9, 2023. doi:10.1007/s00025-023-02010-5.
[15] Philippe Jacquet and Wojciech Szpankowski. Entropy computations via analytic de-Poissonization. *IEEE Trans. Inform. Theory*, 45(4):1072–1081, 1999. doi:10.1109/18.761251.
[16] Daniel Kane and Terence Tao. A bound on partitioning clusters. *Electron. J. Combin.*, 24(2):Paper No. 2.31, 13, 2017. doi:10.37236/6797.
[17] Charles Knessl. Integral representations and asymptotic expansions for Shannon and Renyi entropies. *Appl. Math. Lett.*, 11(2):69–74, 1998. doi:10.1016/S0893-9659(98)00013-5.
[18] Vjekoslav Kovač. On binomial sums, additive energies, and lazy random walks. *J. Math. Anal. Appl.*, 528(1):Paper No. 127510, 13, 2023. doi:10.1016/j.jmaa.2023.127510.
[19] W. Wesley Peterson and E. J. Weldon, Jr. *Error-correcting codes*. The M.I.T. Press, Cambridge, Mass.-London, second edition, 1972.
[20] Tomasz Schoen and Ilya D. Shkredov. Higher moments of convolutions. *J. Number Theory*, 133(5):1693–1737, 2013. doi:10.1016/j.jnt.2012.10.010.
[21] Xuancheng Shao. Additive energies of subsets of discrete cubes. preprint, 2024. URL: https://arxiv.org/abs/2407.06944.
[22] Ilya Shkredov. Energies and structure of additive sets. *Electron. J. Combin.*, 21(3):Paper 3.44, 53, 2014. doi:10.37236/4369.
[23] Terence Tao and Van Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006. doi:10.1017/CBO9780511755149.
[24] Wolfram Research, Inc. Mathematica, Version 13.0.1, 2022.

T. C., Department of Mathematics, Faculty of Science, University of Split, Ruđera Boškovića 33, 21000 Split, Croatia
Email address: tcrmaric@pmfst.hr

V. K., Department of Mathematics, Faculty of Science, University of Zagreb, Bijenička cesta 30, 10000 Zagreb, Croatia
Email address: vjekovac@math.hr