Proceedings of the Royal Society of Edinburgh: Section A Mathematics, 1–27, *2025*
DOI:10.1017/prm.2025.10015

Sharp estimates for Gowers norms on discrete cubes

Adrian Beker
Department of Mathematics, Faculty of Science, University of Zagreb, Bijenička cesta 30, Zagreb, 10000, Croatia (adrian.beker@math.hr)

Tonći Crmarić
Department of Mathematics, Faculty of Science, University of Split, Ruđera Boškovića 33, Split, 21000, Croatia (tcrmaric@pmfst.hr) (corresponding author)

Vjekoslav Kovač
Department of Mathematics, Faculty of Science, University of Zagreb, Bijenička cesta 30, Zagreb, 10000, Croatia (vjekovac@math.hr)

(Received 19 September 2024; revised 10 October 2024; accepted 6 May 2025)

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 $\{0,1,\dots,n-1\}^d$ 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 article:

- 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 $n \geqslant 2$.

Keywords: additive energy; asymptotic formula; binary cube; Lebesgue norm; Shannon entropy

2020 Mathematics Subject Classification: Primary 05D05; Secondary 11B30, 94A17

© The Author(s), 2025. Published by Cambridge University Press on behalf of The Royal Society of Edinburgh

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

1. Introduction

Motivation for this article is twofold. 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

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

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 [26, Section 2.3]. For sets $A \subseteq \{0,1\}^d \subset \mathbb{Z}^d$, Kane and Tao [17, 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 [23] and Shkredov [25]. De Dios Pont, Greenfeld, Ivanisvili, and Madrid [6] revisited two of them and called them the *higher energies*

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

and the *k-additive energies*

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

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 [17, 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 one of the authors of the present article established an ingredient needed for general k

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

in [20]. A third possible quantity that generalizes (1.2) counts (possibly degenerate) parallelotopes with vertices in A, with a lot of overcounting. It is defined as

$$P_k(A) := \big|\big\{(a,h_1,\dots,h_k)\in(\mathbb{Z}^d)^{k+1} \,:\, a+\epsilon_1h_1+\cdots+\epsilon_kh_k\in A \\ \text{for every } (\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k\big\}\big| \tag{1.3}$$

for any $k\geqslant 2$. It has already been studied by Shkredov [25], who also discussed numerous relations between all three kinds of generalized additive energies. It is precisely the generalized energies P_k that are the topic of interest in this article.

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

$$\|f\|_{\mathrm{U}^k} := \Big(\sum_{a,h_1,\dots,h_k}\prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k}\mathcal{C}^{\epsilon_1+\cdots+\epsilon_k}f(a+\epsilon_1h_1+\cdots+\epsilon_kh_k)\Big)^{1/2^k},$$

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 §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 we recover precisely the quantity P_k from (1.3), namely

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

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 [4, Proposition 2.2], or in geometric measure theory of large sets [7, Section 6]. The number $2^k/(k+1)$ was 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

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

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 attempt to control generalized

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

additive energies P_k of $A \subseteq \{0,1\}^d$ by optimal powers of the size of A. Both of these problems are resolved by the first result of this article.

THEOREM 1.1 *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 [24]; also see a more general setting by Hegyvári [14]. In this article, we also want to study non-trivial 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 1.2. *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 1.2 will be borrowed from [6], even though the authors there rather discussed norms of multiple convolutions,

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

and the related k-additive energies E_k. Another difference is that the proofs in [6] use the Fourier transform, while we avoid the need to look at the functions on the

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

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 merely in terms of the size 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 1.2, since part (1) is structurally just an inequality for n non-negative numbers and it can be normalized to yield a compact condition. By the same proposition, we then also know that

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

where $p_{k,n}$ is the largest $p > 0$ such that (1.4) holds in any dimension d and for every function f supported in (1.1).

Theorem 1.1 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, a 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

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

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 [22] based on (1.10) and formula (1.9) gives

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

Inverse Symbolic Calculator [3] does not recognize these two as any of the numbers appearing naturally elsewhere. Still, as a consequence of a general theorem 1.4, 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 behaviour of $t_{k,n}$ as either $n \to \infty$ or $k \to \infty$. Let us first

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

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 [24] 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 1.3 *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 1.3 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 non-trivial bound as a black box. One could, in fact, repeat the proof from [24] 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 article, 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 article. 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 §2 for the more general definition.

THEOREM 1.4 *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 $n \to \infty$;

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

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

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

this is a well-known fact, which also follows from sharper estimates (6.3) in §6. In particular, the constant $t_{k,3}$ relevant for the ternary cube is asymptotically equal to $(4/3)\log_2 k$ and theorem 1.4 also gives the more precise formula (1.11). Somewhat surprisingly, the proof of theorem 1.4 will require subtle estimates for the Shannon entropy of certain independent sums of discrete random variables, which will be discussed in §6.1.

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, $\{1,2,3,\dots\}$. 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. The logarithm with base e will be denoted ln.

If U is a fixed (universal) set, then we write $\mathcal{P}(U)$ for the collection of all of its subsets. We say that two sets $A, B \in \mathcal{P}(U)$ are *comparable* if one of them is a subset of the other. Given a family of subsets $\mathcal{A} \subseteq \mathcal{P}(U)$, we say that $\mathcal{A}$ is a *chain* if any two sets in $\mathcal{A}$ are comparable and that $\mathcal{A}$ is an *antichain* if no two sets in $\mathcal{A}$ are comparable.

Suppose that

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

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

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

if

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

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

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

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

means that

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

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

$$O_{a,b,\dots}^{m\to\infty}\big(G(m,a,b,\dots)\big) \quad \text{and} \quad o_{a,b,\dots}^{m\to\infty}\big(G(m,a,b,\dots)\big)$$

in place of any particular expression $F(m,a,b,\dots)$ 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 asymptotically 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 $p \in [1,\infty]$.

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

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

$$\begin{aligned}&\big\langle (f_{\epsilon_1,\dots,\epsilon_k})_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k}\big\rangle_{\mathrm{U}^k}\\ &:=\sum_{a,h_1,\dots,h_k}\prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k}\mathcal{C}^{\epsilon_1+\dots+\epsilon_k}f_{\epsilon_1,\dots,\epsilon_k}(a+\epsilon_1h_1+\dots+\epsilon_kh_k),\end{aligned}$$

where $(f_{\epsilon_1,\dots,\epsilon_k})$ 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*:

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

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].

It will also be convenient to recall the classical inequality of Karamata [18]. Suppose that we are given two N-tuples

$$\mathbf{x}=(x_1,x_2,\dots,x_N),\quad \mathbf{y}=(y_1,y_2,\dots,y_N)$$

of numbers from an interval $I\subseteq\mathbb{R}$ sorted in the descending order, i.e.,

$$x_1\geqslant x_2\geqslant\dots\geqslant x_N,\quad y_1\geqslant y_2\geqslant\dots\geqslant y_N.$$

We say that $\mathbf{x}$ *majorizes* $\mathbf{y}$ if

$$\sum_{j=1}^{N}x_j=\sum_{j=1}^{N}y_j$$

and

$$\sum_{j=1}^{n}x_j\geqslant\sum_{j=1}^{n}y_j$$

for all $1\leqslant n\leqslant N-1$. Karamata proved that, for a strictly convex function $\psi\colon I\to\mathbb{R}$, we have

$$\sum_{j=1}^{N}\psi(x_j)\geqslant\sum_{j=1}^{N}\psi(y_j)$$

with equality only if $\mathbf{x}=\mathbf{y}$; see the details in [16, Theorem 1, Note 2]. The inequality is reversed if ψ is strictly concave.

Given an integer-valued random variable X on a probability space $(\Omega,\mathcal{F},\mathbb{P})$, we define its *probability mass function* to be the function

$$p_X\colon\mathbb{Z}\to[0,1],\quad p_X(z):=\mathbb{P}(X=z).$$

We will only work with distributions that are supported on finitely many integers. In this case, the *decreasing rearrangement* of p_X is defined to be the decreasing

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

(i.e., non-increasing) function $p_X^{\downarrow}\colon \mathbb{N} \to [0,1]$ taking the same multiset of values as p_X.

The notion of the Shannon entropy will be needed in §6. If a random variable X has distribution supported on a finite subset of $\mathbb{Z}$, then its *Shannon entropy* (or simply just *entropy*) is defined to be the number

$$\mathrm{H}(X) := -\sum_{z\in\mathbb{Z}} p_X(z)\log_2 p_X(z). \tag{2.2}$$

Here we interpret $0\log_2 0$ as 0. In the particular case of a random variable X distributed

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

its entropy will sometimes also be written as

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

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

3. Proof of proposition 1.2

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

$$pt = 2^k.$$

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

$$\sum_{b,l_1,\dots,l_k\in\mathbb{Z}} \prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k} g(b+\epsilon_1 l_1+\cdots+\epsilon_k l_k) \leqslant \Big(\sum_{b=0}^{n-1} g(b)^p\Big)^{2^k/p}. \tag{3.1}$$

We prove the claim (2) by induction on d. The base 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

$$\|f_b\|_{\mathrm{U}^k} \leqslant \|f_b\|_{\ell^p}. \tag{3.2}$$

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

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

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

so the Gowers–Cauchy–Schwarz inequality followed by (3.2) 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 (3.1) 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) $\implies$ (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) $\implies$ (4). This is obvious by taking $C = 1$.

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

$$\|\mathbb{1}_A\|_{\mathrm{U}^k} \leqslant D|A|^{t/2^k} = D\|\mathbb{1}_A\|_{\ell^p} \tag{3.3}$$

for every $d \geqslant 0$ and every $A \subseteq \{0,1,\dots,n-1\}^d$. 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

$$\|f\|_{\mathrm{U}^k} \leqslant D(3 + d\log_2 n)\|f\|_{\ell^p}. \tag{3.4}$$

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

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

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

for every $a \in \{0,1,\dots,n-1\}^d$ 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 (3.3) 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

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

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 (3.4).

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

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

Then

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

so taking the dth roots of (3.4) 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).

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

4. Proof of theorem 1.1

Fix an integer $k \geqslant 2$ throughout this section and set

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

always remembering that t depends on k. By proposition 1.2, i.e., by formula (1.9), we only need to verify

$$\sum_{a,h_1,\dots,h_k \in \mathbb{Z}} \prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k} f(a+\epsilon_1 h_1+\cdots+\epsilon_k h_k) \leqslant \left(f(0)^{2^k/t}+f(1)^{2^k/t}\right)^t \quad (4.1)$$

for every $f\colon \mathbb{Z}\to[0,\infty)$ that vanishes outside $\{0,1\}$. Substituting

$$x = f(0)^{2^k/t}, \quad y = f(1)^{2^k/t}$$

we evaluate the left-hand side of (4.1) as follows.

- The term with $a=0$, $h_1=\cdots=h_k=0$ contributes to the sum with $f(0)^{2^k}=x^t$.
- The term with $a=1$, $h_1=\cdots=h_k=0$ contributes to the sum with $f(1)^{2^k}=y^t$.
- Terms with $a=0$, $h_{j_0}=1$ for some index j_0, $h_j=0$ for every index $j\neq j_0$ contribute to the sum with $kf(0)^{2^k-1}f(1)^{2^{k-1}}=kx^{t/2}y^{t/2}$.
- Terms with $a=1$, $h_{j_0}=-1$ for some index j_0, $h_j=0$ for every index $j\neq j_0$ also contribute to the sum with $kf(0)^{2^k-1}f(1)^{2^{k-1}}=kx^{t/2}y^{t/2}$.

It remains to prove the generalization of an 'elementary inequality' by Kane and Tao [17, Lemma 8] stated in lemma 4.1. It is similar to but different from [6, Lemma 9], [6, Lemma 5], and [20, Theorem 1].

LEMMA 4.1. *For every $x,y\in[0,\infty)$ we have*

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

Proof. The case $k=2$ is precisely [17, Lemma 8], which has also been reproved several times in [6, 20]. 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

$$\varphi(x) \leqslant 1 \quad (4.2)$$

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

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 (4.2) 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 (4.2).

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

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

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 (4.2) holds on that interval too. □

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

5. Proof of theorem 1.3

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

$$\|f\|_{\mathrm{U}^k(\mathbb{R})}:=\Bigg(\int_{\mathbb{R}^{k+1}}\prod_{(\epsilon_1,\dots,\epsilon_k)\in\{0,1\}^k}\times\mathcal{C}^{\epsilon_1+\cdots+\epsilon_k}f(x+\epsilon_1h_1+\cdots+\epsilon_kh_k)\,\mathrm{d}x\,\mathrm{d}h_1\cdots\mathrm{d}h_k\Bigg)^{1/2^k}.$$

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 [24], 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.

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

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

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

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

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

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

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

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),

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

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

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

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

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 side 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 induction on $k\geqslant 2$. The base $k=2$ 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 1.4

Before we begin with the proof of theorem 1.4, 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

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

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

6.1. Entropy estimates

Let $X_1, X_2, X_3, \dots$ be independent symmetric Bernoulli random variables, namely

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

so that $X_1 + X_2 + \cdots + X_m \sim B(m, 1/2)$. This subsection will be concerned with random variables of the form

$$h_1 X_1 + \cdots + h_m X_m$$

for arbitrary non-zero integers $h_1, \dots, h_m$.

First, we need precise estimates for 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 [11, 15, 19], but we need exact inequalities. They are available from the work of Adell, Lekuona, and Yu [1], who strengthened the previous bounds by Chang and Weldon [5].

LEMMA 6.1. [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] used the natural logarithm in the definition of the entropy (i.e., they measure it in nats), whereas we chose the base-2-logarithm in (2.2) (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

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

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. □

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

COROLLARY 6.2. *The sequence* $(H_m/m)_{m=1}^{\infty}$ *is strictly decreasing, i.e.,*

$$\frac{H_1}{1} > \frac{H_2}{2} > \frac{H_3}{3} > \frac{H_4}{4} > \cdots .$$

Proof. For any integer $m \geqslant 2$ inequality

$$(m+1)H_m > mH_{m+1}$$

will follow from lemma 6.1 and

$$(m+1)\Big(\frac{1}{2}\log_2 \frac{e\pi m}{2} - \frac{1}{4m}\Big) > m\Big(\frac{1}{2}\log_2 \frac{e\pi(m+1)}{2} + \frac{1}{10(m+1)}\Big).$$

However, the last estimate can be multiplied by 2 and equivalently rewritten as

$$\underbrace{\log_2 m}_{\geqslant 1} + \underbrace{\log_2 \frac{\pi}{2}}_{>0} + \underbrace{\log_2 e - \log_2 \Big(1+\frac{1}{m}\Big)^m}_{>0} > \underbrace{\frac{m}{5(m+1)} + \frac{m+1}{2m}}_{<1/5+3/4<1},$$

which clearly holds. Finally, the remaining inequality between the first two binomial entropies is trivial because of $H_1 = 1$, $H_2 = 3/2$. □

We turn to formulating and proving more general inequalities for the Shannon entropy that will play a crucial role in the proof of theorem 1.4. Recall that p_Z denotes the probability mass function of a discrete, integer-valued random variable Z; see §2. Also recall that $p_Z^{\downarrow}$ is its decreasing rearrangement.

Let us begin with a purely combinatorial result.

LEMMA 6.3. *Let* $h_1, \ldots, h_m$ *be non-zero integers,* $X = h_1X_1 + \ldots + h_mX_m$*, and* $Y \sim B(m, 1/2)$*. Then* $p_Y^{\downarrow}$ *majorizes* $p_X^{\downarrow}$*. Moreover, we have* $p_X^{\downarrow} = p_Y^{\downarrow}$ *if and only if* $|h_1| = \cdots = |h_m|$*.*

Here we view $p_X^{\downarrow}$ and $p_Y^{\downarrow}$ as N-tuples for a positive integer N sufficiently large that both of their supports are contained in $\{1, 2, \ldots, N\}$. In this situation, the notion of 'majorization' from §2 applies too.

Proof. Changing some signs among $h_1, \ldots, h_m$ only translates the distribution of $h_1X_1 + \cdots + h_mX_m$ 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. Also, the case $m = 1$ is trivial because of $X \sim Y$, so we assume that $m \geqslant 2$.

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

Consider, for each $z \in \mathbb{Z}$, the family

$$\mathcal{A}_z := \Big\{ A \subseteq \{1, 2, \dots, m\} \,:\, \sum_{i \in A} h_i = z \Big\},$$

which is readily seen to be an antichain. It is clear that

$$\sum_{j=1}^{\infty} p_Y^{\downarrow}(j) = 1 = \sum_{j=1}^{\infty} p_X^{\downarrow}(j).$$

Furthermore, given any $n \in \mathbb{N}$, we want to show that

$$\sum_{j=1}^{n} p_Y^{\downarrow}(j) \geqslant \sum_{j=1}^{n} p_X^{\downarrow}(j). \tag{6.4}$$

On the one hand, observe that

$$p_Y^{\downarrow}(1),\ p_Y^{\downarrow}(2),\ p_Y^{\downarrow}(3),\ p_Y^{\downarrow}(4),\ \dots$$

appearing on the left-hand side of (6.4) are simply $(1/2^m)$-multiples of the binomial coefficients with m on top, sorted in the descending order; namely,

$$2^{-m}\binom{m}{\lfloor m/2 \rfloor},\ 2^{-m}\binom{m}{\lfloor m/2 \rfloor + 1},\ 2^{-m}\binom{m}{\lfloor m/2 \rfloor - 1},\ 2^{-m}\binom{m}{\lfloor m/2 \rfloor + 2},\ \dots.$$

On the other hand, choose a bijection $\pi\colon \mathbb{N} \to \mathbb{Z}$ such that $p_X^{\downarrow} = p_X \circ \pi$. The right-hand side of (6.4) equals $|\mathcal{B}_n|/2^m$, where we define

$$\mathcal{B}_n := \bigcup_{j=1}^{n} \mathcal{A}_{\pi(j)}.$$

Note that $\mathcal{B}_n$ is a union of n antichains in $\mathcal{P}(\{1, 2, \dots, m\})$, so, in particular, any chain in $\mathcal{B}_n$ has size at most n. Erdős [9] proved a generalization of Sperner's theorem bounding the maximal size of a subcollection of $\mathcal{P}(\{1, 2, \dots, m\})$ that does not have a chain of length longer than n. Namely, by [9, Theorem 5], it follows that $|\mathcal{B}_n|$ is at most the sum of the n largest binomial coefficients of the form $\binom{m}{k}$ with $0 \leqslant k \leqslant m$. Therefore, the inequality (6.4) follows.

Turning to the second statement, it is clear that $h_1 = \cdots = h_m$ implies $p_X^{\downarrow} = p_Y^{\downarrow}$. To prove the converse, note that $p_X^{\downarrow}(1) = p_Y^{\downarrow}(1)$ implies that $|\mathcal{A}_{\pi(1)}| = \binom{m}{\lfloor m/2 \rfloor}$. Since $\mathcal{A}_{\pi(1)}$ is an antichain, it follows from the equality case in Sperner's theorem (see, e.g., [2, Theorem 1.2.2]) that $\mathcal{A}_{\pi(1)}$ is the family of all subsets of $\{1, 2, \dots, m\}$ of some fixed size $k \in \{\lfloor m/2 \rfloor, \lceil m/2 \rceil\}$. But now, for any distinct $i, j \in \{1, \dots, m\}$, we may choose a subset $A \subseteq \{1, \dots, m\} \setminus \{i, j\}$ of size $k - 1$. In particular, $A \cup \{i\}$ and $A \cup \{j\}$ both belong to $\mathcal{A}_{\pi(1)}$. This means that

$$h_i + \sum_{a \in A} h_a = \sum_{a \in A \cup \{i\}} h_a = \pi(1) = \sum_{a \in A \cup \{j\}} h_a = h_j + \sum_{a \in A} h_a,$$

whence $h_i = h_j$, as desired. □

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

REMARK 6.4. The fact that the size of a union of n antichains in $\mathcal{P}(\{1,2,\dots,m\})$ is maximized by taking the n largest layers can also be easily proved by means of the Yamamoto–Bollobás–Lubell–Meshalkin inequality [21].

The following consequence of the previous lemma can be thought of as an entropic variant of the Littlewood–Offord problem [9].

PROPOSITION 6.5. *For every $m \in \mathbb{N}$ and arbitrary $h_1, \dots, h_m \in \mathbb{Z} \setminus \{0\}$ we have*

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

with equality attained only when $|h_1| = \dots = |h_m|$.

Note that the case of equality, namely $|h_1| = \dots = |h_m|$, means precisely that $h_1X_1 + \dots + h_mX_m$ is distributed as a non-degenerate affine transformation of the symmetric binomial distribution $B(m, 1/2)$.

Proof. Consider the function

$$\psi\colon [0,1] \to \mathbb{R}, \quad \psi(x) := \begin{cases} -x \log_2 x & \text{for } x > 0, \\ 0 & \text{for } x = 0. \end{cases}$$

Then the entropy (2.2) of an arbitrary integer-valued random variable Z taking only finitely many values can be expressed as

$$H(Z) = \sum_{z \in \mathbb{Z}} \psi(p_Z(z)) = \sum_{n=1}^{\infty} \psi(p_Z^{\downarrow}(n)).$$

Again take $X = h_1X_1 + \dots + h_mX_m$ and $Y \sim B(m, 1/2)$. From lemma 6.3, we know that $p_Y^{\downarrow}$ majorizes $p_X^{\downarrow}$. Since ψ is strictly concave, the desired estimate,

$$H(X) \geqslant H(Y),$$

follows from Karamata's inequality recalled in §2. Equality holds if and only if $p_X^{\downarrow} = p_Y^{\downarrow}$, which is, by lemma 6.3 again, satisfied only when all h_i are the same in the absolute value. □

COROLLARY 6.6. *For every $n \geqslant 2$, $1 \leqslant l \leqslant n-1$, and $h_1, \dots, h_l \in \mathbb{Z} \setminus \{0\}$ such that*

$$|h_1| + \dots + |h_l| \leqslant n - 1 \tag{6.5}$$

we have

$$\frac{\mathrm{H}(h_1X_1 + \dots + h_lX_l)}{l} \geqslant \frac{H_{n-1}}{n-1}, \tag{6.6}$$

with equality attained only when $l = n-1$ and $|h_1| = \dots = |h_{n-1}| = 1$, in which case $h_1X_1 + \dots + h_{n-1}X_{n-1}$ is distributed as a translate of $B(n-1, 1/2)$.

Proof. If the number of terms l is fixed, then, by proposition 6.5, the left-hand side of (6.6) is at least H_l/l and then the strict inequality for $1 \leqslant l \leqslant n-2$ follows

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

from corollary 6.2. The inequality also holds when $l = n-1$ and, by proposition 6.5 again, we must additionally have $|h_1| = \cdots = |h_{n-1}|$ in the case of an equality. This combined with the assumption (6.5) means that all these absolute values need to be 1. □

Remark 6.7. Corollary 6.6 will be sufficient for our intended application and it allows for a conceptually simpler, but more computational proof. Namely, the most classical result on the Littlewood–Offord problem [9, Theorem 1],

$$\max_{z\in\mathbb{Z}} \mathbb{P}(h_1X_1 + \cdots + h_lX_l = z) \leqslant \binom{l}{\lfloor l/2\rfloor} 2^{-l},$$

gives a 'cheaper' lower bound

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

For $n > 100$ and $1 \leqslant l \leqslant 3(n-1)/4$, we then use lemma 6.1 and some elementary calculus to show

$$\begin{aligned}
&2(n-1)\mathrm{H}(h_1X_1 + \cdots + h_lX_l) - 2lH_{n-1}\\
&> (n-1)\log_2\frac{\pi l}{2} - 2lH_{n-1}\\
&> (n-1)\log_2\frac{\pi l}{2} - l\log_2\frac{e\pi(n-1)}{2} - \frac{l}{5(n-1)} > 0.
\end{aligned}$$

On the other hand, when $n > 100$ and $3(n-1)/4 < l \leqslant n-2$, the assumption (6.5) forces at least $2l - n + 1$ of the numbers h_i to be ± 1, which easily gives

$$\mathrm{H}(h_1X_1 + \cdots + h_lX_l) \geqslant H_{2l-n+1}.$$

Then lemma 6.1 and basic calculus yield

$$\begin{aligned}
&2(n-1)\mathrm{H}(h_1X_1 + \cdots + h_lX_l) - 2lH_{n-1}\\
&\geqslant 2(n-1)H_{2l-n+1} - 2lH_{n-1}\\
&> (n-1)\log_2\frac{e\pi(2l-n+1)}{2} - \frac{n-1}{2(2l-n+1)} - l\log_2\frac{e\pi(n-1)}{2} - \frac{l}{5(n-1)} > 0.
\end{aligned}$$

Finally, if $n \leqslant 100$, then (6.5) leaves only finitely many cases to check, and a combination of the previous reasoning and a computer-assisted verification handles all of them.

6.2. The lower bound

Recall that, by proposition 1.2, the number $t_{k,n}$ is the smallest $t > 0$ such that

$$\sum_{a,h_1,\ldots,h_k\in\mathbb{Z}} \prod_{(\epsilon_1,\ldots,\epsilon_k)\in\{0,1\}^k} f(a + \epsilon_1h_1 + \cdots + \epsilon_kh_k) \leqslant \left(\sum_{j=0}^{n-1} f(j)^{2^k/t}\right)^t \tag{6.7}$$

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

holds for every function $f\colon \mathbb{Z} \to [0,\infty)$ supported in $\{0,1,\dots,n-1\}$. On the left-hand side of inequality (6.7), we only observe the mutually equal terms obtained by taking:

- $a \in \{0,1,\dots,n-1\}$ arbitrary,
- precisely a of the numbers $h_1,\dots,h_k$ to be equal to -1,
- precisely $n-1-a$ of the numbers $h_1,\dots,h_k$ to be equal to 1, and
- precisely $k-n+1$ of the numbers $h_1,\dots,h_k$ to be equal to 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 (6.7) with the same term

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

Thus, by using inequality (6.7) with $t = t_{k,n}$, we obtain

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

Now take a particular function f defined as

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

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

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

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.,

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

which proves (6.1).

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

6.3. The upper bound

Note that the term on the left-hand side of (6.7) corresponding to at least n non-zero numbers $h_1, \dots, h_k$ 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 $1 \leqslant l \leqslant n-1$ non-zero numbers among $h_1, \dots, h_k$ in $\binom{k}{l}$ ways. Therefore, for a fixed integer $n \geqslant 2$ and every $l \in \{1, \dots, n-1\}$, we define

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

so that inequality (6.7) can now be rewritten as

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

By substituting $g(j) = f(j)^{2^k/t}$, the estimate (6.7) is further equivalent to

$$\sum_{j=0}^{n-1} g(j)^t + \sum_{l=1}^{n-1} \sum_{(a,h_1,\dots,h_l)\in T_{n,l}} \binom{k}{l} \prod_{(\epsilon_1,\dots,\epsilon_l)\in\{0,1\}^l} g(a+\epsilon_1 h_1 + \cdots + \epsilon_l h_l)^{t/2^l} \leqslant 1 \quad (6.8)$$

for every $g\colon \{0, 1, \dots, n-1\} \to [0, \infty)$ 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 (6.8) for any given $0 < \delta < 1$, for

$$t = \frac{(n-1)\log_2(2k) - \log_2(n-1)!}{H_{n-1}} + \delta, \quad (6.9)$$

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 (6.8) is of the form

$$\binom{k}{l} \big(g(0)^{q_0} g(1)^{q_1} \cdots g(n-1)^{q_{n-1}}\big)^t \quad (6.10)$$

where $(q_0, q_1, \dots, q_{n-1})$ is some discrete probability distribution on the set $\{0, 1, \dots, n-1\}$, which is also the distribution of some random variable $a + h_1 X_1 + \cdots + h_l X_l$ for $(a, h_1, \dots, h_l) \in T_{n,l}$. We always interpret $0^0 = 1$. For every g as before the inequality between the weighted arithmetic and geometric means gives

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

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

That way we have obtained

$$\prod_{j=0}^{n-1} g(j)^{q_j} \leqslant 2^{-\mathrm{H}(q_0,\dots,q_{n-1})}, \tag{6.11}$$

where we recall that $\mathrm{H}(q_0,\dots,q_{n-1})$ is defined by the formula (2.2). Denote

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

The proof of (6.8) 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 (6.10) from the left-hand side of (6.8) corresponding to some $(a, h_1, \dots, h_l) \in T_{n,l}$ is, by inequality (6.11), at most

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

which is, by the choice (6.9) for t, less than or equal to

$$\begin{cases} 2^{-n+1}\cdot 2^{-H_{n-1}\delta} & \text{for } l = n-1, \qquad (6.12)\\ O_n^{k\to\infty}\big(k^{l-(n-1)\mathrm{H}(h_1X_1+\dots+h_lX_l)/H_{n-1}}\big) & \text{for } 1\leqslant l\leqslant n-2. \qquad (6.13)\end{cases}$$

In (6.12), we used the fact that the only possibility to have $l = n-1$ is $a+h_1X_1+\dots+h_{n-1}X_{n-1} \sim B(n-1,1/2)$. Since $T_{n,n-1}$ has precisely 2^{n-1} elements described in the previous subsection and each corresponding term is bounded by (6.12), we conclude

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

Next, each term corresponding to $T_{n,l}$, $1\leqslant l\leqslant n-2$, is bounded by (6.13) and the exponent of k is strictly negative by corollary 6.6. Thus,

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

Finally, under the standing assumptions on g and by formula (6.9),

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

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

Altogether, the left-hand side of (6.8) 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 so that $t > 2^n$; a further largeness requirement on k will be imposed later. The terms $g(j)^t$ in (6.8) 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 (6.8) is at most

$$k^{n-1} g(j_1)^{t/2^n} g(j_2)^{t/2^n} \tag{6.14}$$

for some distinct indices j_1 and j_2. Since at least one of these indices is different from j_0, we can bound (6.14) 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 (6.8) 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 (6.8), and thus also of the upper bound (6.2).

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

References

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

[2] I. Anderson. *Combinatorics of finite sets* (Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1987).

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press

[3] P. B. Borwein, J. M. Borwein, S. Plouffe. Inverse symbolic calculator, 1995. https://wayback.cecm.sfu.ca/projects/ISC/ISCmain.html

[4] J. Byszewski, J. Konieczny, and C. Müllner. Gowers norms for automatic sequences. *Discrete Anal* **4** (2023), 62.

[5] S. C. Chang and E. J. Weldon. Coding for T-user multiple-access channels. *IEEE Trans. Inform. Theory* **25** (1979), 684–691. https://doi.org/10.1109/TIT.1979.1056109

[6] J. de Dios Pont, R. Greenfeld, P. Ivanisvili and J. Madrid. Additive energies on discrete cubes. *Discrete Anal.* **13** (2023), 16. https://doi.org/10.19086/da.84737

[7] P. Durcik and V. Kovač. A Szemerédi-type theorem for subsets of the unit cube. *Anal. PDE* **15** (2022), 507–549. https://doi.org/10.2140/apde.2022.15.507

[8] T. Eisner and T. Tao. Large values of the Gowers-Host-Kra seminorms. *J. Anal. Math.* **117** (2012), 133–186. https://doi.org/10.1007/s11854-012-0018-2

[9] P. Erdos. On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc.* **51** (1945), 898–902. https://doi.org/10.1090/S0002-9904-1945-08454-7

[10] A. Fan and J. Konieczny. On uniformity of q-multiplicative sequences. *Bull. Lond. Math. Soc.* **51** (2019), 466–488. https://doi.org/10.1112/blms.12245

[11] P. Flajolet. Singularity analysis and asymptotics of Bernoulli sums. *Theoret. Comput. Sci.* **215** (1999) 371–381. https://doi.org/10.1016/S0304-39759800220-5.

[12] W. T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.* **8** (1998), 529–551. https://doi.org/10.1007/s000390050065

[13] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.* **11** (2001), 465–588. https://doi.org/10.1007/s00039-001-0332-9

[14] N. Hegyvári. Some remarks on the distribution of additive energy. *Results Math.* **78** (2023), 233, 9. https://doi.org/10.1007/s00025-023-02010-5

[15] P. Jacquet and W. Szpankowski. Entropy computations via analytic de-Poissonization. *IEEE Trans. Inform. Theory* **45** (1999), 1072–1081. https://doi.org/10.1109/18.761251

[16] Z. Kadelburg, D. Djukić, M. Lukić, and I. Matić. Inequalities of Karamata, Schur and Muirhead, and some applications. *The Teaching of Mathematics* **8** (2005), 31–45.

[17] D. Kane and T. Tao. A bound on partitioning clusters. *Electron. J. Combin.* **24** (2017), 2.31, 13. https://doi.org/10.37236/6797

[18] J. Karamata. Sur une inégalité relative aux fonctions convexes. *Publ. Math. Univ. Belgrade* **1** (1932), 145–148

[19] C. Knessl. Integral representations and asymptotic expansions for Shannon and Renyi entropies. *Appl. Math. Lett.* **9659** (1998), 69–74. https://doi.org/10.1016/S0893-00013-5.

[20] V. Kovač. On binomial sums, additive energies, and lazy random walks. *J. Math. Anal. Appl.* **528** (2023), 127510, 13. https://doi.org/10.1016/j.jmaa.2023.127510

[21] D. Lubell. A short proof of Sperner's lemma. *J. Combinatorial Theory* **1** (1966), 299.

[22] Wolfram Research Inc. Mathematica, Version 13.0.1 (2022).

[23] T. Schoen and D. S. Ilya. Higher moments of convolutions. *J. Number Theory* **133** (2013), 1693–1737. https://doi.org/10.1016/j.jnt.2012.10.010

[24] X. Shao. Additive energies of subsets of discrete cubes. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* (2024), 1–22. https://doi.org/10.1017/prm.2024.126

[25] I. Shkredov. Energies and structure of additive sets. *Electron. J. Combin.* **21** (2014), 3.44, 53. https://doi.org/10.37236/4369

[26] T. Tao and V. Van. *Additive combinatorics*. Of Cambridge Studies in Advanced Mathematics, Vol. 105 (Cambridge University Press, Cambridge, 2006). https://doi.org/10.1017/CBO9780511755149

https://doi.org/10.1017/prm.2025.10015 Published online by Cambridge University Press