

ARTICLE

# Approximate discrete entropy monotonicity for log-concave sums

Lampros Gavalakis 

Laboratoire d'Analyse et de Mathématiques Appliquées, Université Gustave Eiffel, Champs-sur-Marne, France  
Email: lampros.gavalakis@univ-eiffel.fr

(Received 26 November 2022; revised 4 October 2023; accepted 16 October 2023; first published online 13 November 2023)

## Abstract

It is proven that a conjecture of Tao (2010) holds true for log-concave random variables on the integers: For every  $n \geq 1$ , if  $X_1, \dots, X_n$  are i.i.d. integer-valued, log-concave random variables, then

$$H(X_1 + \dots + X_{n+1}) \geq H(X_1 + \dots + X_n) + \frac{1}{2} \log \left( \frac{n+1}{n} \right) - o(1)$$

as  $H(X_1) \rightarrow \infty$ , where  $H(X_1)$  denotes the (discrete) Shannon entropy. The problem is reduced to the continuous setting by showing that if  $U_1, \dots, U_n$  are independent continuous uniforms on  $(0, 1)$ , then

$$h(X_1 + \dots + X_n + U_1 + \dots + U_n) = H(X_1 + \dots + X_n) + o(1),$$

as  $H(X_1) \rightarrow \infty$ , where  $h$  stands for the differential entropy. Explicit bounds for the  $o(1)$ -terms are provided.

**Keywords:** Entropy; monotonicity; log-concavity; entropy power inequality; central limit theorem; concentration; sunset inequalities

**2020 MSC Codes:** Primary: 94A17; Secondary: 60E15, 39B62

## 1. Introduction

### 1.1 Monotonic increase of differential entropy

Let  $X, Y$  be two independent random variables with densities in  $\mathbb{R}$ . The differential entropy of  $X$ , having density  $f$ , is

$$h(X) = - \int_{\mathbb{R}} f(x) \log f(x) dx$$

and similarly for  $Y$ . Throughout ‘log’ denotes the natural logarithm.

The entropy power inequality (EPI) plays a central role in information theory. It goes back to Shannon [15] and was first proven in full generality by Stam [16]. It asserts that

$$N(X + Y) \geq N(X) + N(Y), \tag{1}$$

where  $N(X)$  is the *entropy power* of  $X$ :

$$N(X) = \frac{1}{2\pi e} e^{2h(X)}.$$

---

L.G. has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 101034255. 

If  $X_1, X_2$  are identically distributed, (1) can be rewritten as

$$h(X_1 + X_2) \geq h(X_1) + \frac{1}{2} \log 2. \tag{2}$$

The EPI is also connected with and has applications in probability theory. The following generalisation is due to Artstein, Ball, Barthe and Naor [1]: If  $\{X_i\}_{i=1}^{n+1}$  are continuous, i.i.d. random variables then

$$h\left(\frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} X_i\right) \geq h\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i\right). \tag{3}$$

This is the monotonic increase of entropy along the central limit theorem [2]. The main result of this paper may be seen as an approximate, discrete analogue of (3).

### 1.2 Sumset theory for entropy

There has been interest in formulating discrete analogues of the EPI from various perspectives [7–9, 20]. It is not hard to see that the exact statement (2) can not hold for all discrete random variables by considering deterministic (or even close to deterministic) random variables.

Suppose  $G$  is an additive abelian group and  $X$  is a random variable supported on a discrete (finite or countable) subset  $A$  of  $G$  with probability mass function (p.m.f.)  $p$  on  $G$ . The Shannon entropy, or simply entropy of  $X$  is

$$H(X) = - \sum_{x \in A} p(x) \log p(x). \tag{4}$$

Tao [17] proved that if  $G$  is torsion-free and  $X$  takes finitely many values then

$$H(X_1 + X_2) \geq H(X_1) + \frac{1}{2} \log 2 - o(1), \tag{5}$$

where  $X_1, X_2$  are independent copies of  $X$  and the  $o(1)$ -term vanishes as the entropy of  $X$  tends to infinity. That work explores the connection between additive combinatorics and entropy, which was identified by Tao and Vu in the unpublished notes [18] and by Ruzsa [14]. The main idea is that random variables in  $G$  may be associated with subsets  $A$  of  $G$ : By the *asymptotic equipartition property* [4], there is a set  $A_n$  (the typical set) such that if  $X_1, \dots, X_n$  are i.i.d. copies of  $X$ , then  $(X_1, \dots, X_n)$  is approximately uniformly distributed on  $A_n$  and  $|A_n| = e^{n(H(X)+o(1))}$ . Hence, given an inequality involving cardinalities of sumsets, it is natural to guess that a counterpart statement holds true for random variables if the logarithm of the cardinality is replaced by the entropy.

Exploring this connection, Tao [17] proved an inverse theorem for entropy, which characterises random variables for which the addition of an independent copy does not increase the entropy by much. This is the entropic analogue of the inverse Freiman theorem [19] from additive combinatorics, which characterises sets for which the sumset is not much bigger than the set itself. The discrete EPI (5) is a consequence of the inverse theorem for entropy.

Furthermore, it was conjectured in [17] that for any  $n \geq 2$  and  $\epsilon > 0$

$$H(X_1 + \dots + X_{n+1}) \geq H(X_1 + \dots + X_n) + \frac{1}{2} \log \left(\frac{n+1}{n}\right) - \epsilon, \tag{6}$$

provided that  $H(X)$  is large enough depending on  $n$  and  $\epsilon$ , where  $\{X_i\}_{i=1}^{n+1}$  are i.i.d. copies of  $X$ .

We will prove that the conjecture (6) holds true for log-concave random variables on the integers. An important step in the proof of (5) is reduction to the continuous setting by approximation of the continuous density with a discrete p.m.f.; we briefly outline these key points from that proof in Section 1.3 below as we are going to take a similar approach.

A discrete entropic central limit theorem was recently established in [6]. A discussion relating the above conjecture to the convergence of Shannon entropy to its maximum in analogy with (3) may be found there.

It has also been of interest to establish finite bounds for the  $o(1)$ -term [7, 20]. Our proofs yield explicit rates for the  $o(1)$ -terms, which are exponential in  $H(X_1)$ .

The class of discrete log-concave distributions has been considered recently by Bobkov, Marsiglietti and Melbourne [3] in connection with the EPI. In particular, discrete analogues of (1) were proved for this class. In addition, sharp upper and lower bounds on the maximum probability of discrete log-concave random variables in terms of their variance were provided, which we are going to use in the proofs (see Lemma 5 below). Although log-concavity is a strong assumption in that it implies, for example, connected support set and moments of all orders, many important distributions are log-concave, e.g. Bernoulli, Poisson, geometric, negative binomial and others.

**1.3 Main results and proof ideas**

The first step in the proof method of [17, Theorem 1.9] is to assume that  $H(X_1 + X_2) \leq H(X_1) + \frac{1}{2} \log 2 - \epsilon$ . Then, because of [17, Theorem 1.8], proving the result for random variables  $X$  that can be expressed as a sum  $Z + U$ , where  $Z$  is a random variable with entropy  $O(1)$  and  $U$  is a uniform on a large arithmetic progression, say  $P$ , suffices to get a contradiction. Such random variables satisfy, for every  $x$ ,

$$\mathbb{P}(X = x) \leq \frac{C}{|P|} \tag{7}$$

for some absolute constant  $C$ . Using tools from the theory of sum sets, it is shown that it suffices to consider random variables that take values in a finite subset of the integers. For such random variables that satisfy (7), the smoothness property

$$\|p_{X_1+X_2} - p_{X_1+X_2+1}\|_{TV} \rightarrow 0 \tag{8}$$

as  $H(X) \rightarrow \infty$  is established, where  $p_{X_1+X_2}, p_{X_1+X_2+1}$  are the p.m.f.s of  $X_1 + X_2$  and  $X_1 + X_2 + 1$ , respectively, and  $\|\cdot\|_{TV}$  is the total variation distance defined in (15) below. Using this, it is shown that

$$h(X_1 + X_2 + U_1 + U_2) = H(X_1 + X_2) + o(1), \tag{9}$$

as  $H(X) \rightarrow \infty$ , where  $U_1, U_2$  are independent continuous uniforms on  $(0, 1)$ . The EPI for continuous random variables is then invoked.

The tools that we use are rather probabilistic—our proofs lack any additive combinatorial arguments as we already work with random variables on the integers, which have connected support set. An important technical step in our case is to show that any log-concave random variable  $X$  on the integers satisfies

$$\|p_X - p_{X+1}\|_{TV} \rightarrow 0$$

as  $H(X) \rightarrow \infty$ . Using this we show a generalisation of (9), our main technical tool:

**Theorem 1.** *Let  $n \geq 1$  and suppose  $X_1, \dots, X_n$  are i.i.d. log-concave random variables on the integers with common variance  $\sigma^2$ . Let  $U_1, \dots, U_n$  be continuous i.i.d. uniforms on  $(0, 1)$ . Then*

$$h(X_1 + \dots + X_n + U_1 + \dots + U_n) = H(X_1 + \dots + X_n) + o(1), \tag{10}$$

where the  $o(1)$ -term vanishes as  $\sigma^2 \rightarrow \infty$  depending on  $n$ . In fact, this term can be bounded absolutely by

$$2^{n+6} e^{-(\sqrt{n}\sigma)^{1/5}} (\sqrt{n}\sigma)^3 + \frac{2^{n+2}}{\sigma\sqrt{n}} \log(2^{n+2}\sigma\sqrt{n}) + \frac{\log(n\sigma^2)}{8n\sigma^2}, \tag{11}$$

provided that  $\sigma > \max\{2^{n+2}/\sqrt{n}, 3^7/\sqrt{n}\}$ .

**Remark 2.** Note that always,  $H(X) \rightarrow \infty$  implies  $\sigma^2 \rightarrow \infty$ , since by the maximum entropy property of the Gaussian distribution [4]

$$H(X) = h(X + U) \leq \frac{1}{2} \log \left( 2\pi e\sigma^2 \left( 1 + \frac{1}{12} \right) \right), \tag{12}$$

where  $U$  is an independent continuous uniform on  $(0, 1)$ . Conversely, for the class of log-concave random variables  $\sigma \rightarrow \infty$  implies  $H(X) \rightarrow \infty$ , e.g. by Proposition 7 and Lemma 5. Indeed these give a quantitative comparison between  $H(X)$  and  $\sigma^2 = \text{Var}(X)$  for log-concave random variables:  $H(X) \geq \log \sigma$  for  $\sigma \geq 1$ .

Our main tools are first, to approximate the density of the log-concave sum convolved with the sum of  $n$  continuous uniforms with the discrete p.m.f. (Lemma 8) and second, to show a type of concentration for the ‘‘information density’’,  $-\log p(S_n)$ , using Lemma 9. It is a standard argument to show that log-concave p.m.f.s have exponential tails, since the sum of the probabilities is convergent. Lemma 9 is a slight improvement in that it provides a bound for the ratio depending on the variance.

By an application of the generalised EPI for continuous random variables, we show that the conjecture (6) is true for log-concave random variables on the integers, with an explicit dependence between  $H(X)$  and  $\epsilon$ . Our main result is:

**Theorem 3.** *Let  $n \geq 1$  and  $\epsilon \in (0, 1)$ . Suppose  $X_1, \dots, X_n$  are i.i.d. log-concave random variables on the integers. Then if  $H(X_1)$  is sufficiently large depending on  $n$  and  $\epsilon$ ,*

$$H(X_1 + \dots + X_{n+1}) \geq H(X_1 + \dots + X_n) + \frac{1}{2} \log \left( \frac{n+1}{n} \right) - \epsilon. \tag{13}$$

In fact, for (13) to hold it suffices to take  $H(X_1) \geq \log \frac{2}{\epsilon} + \log \log \frac{2}{\epsilon} + n + 27$ .

The proofs of Theorems 1 and 3 are given in Section 3. Before that, in Section 2 below, we prove some preliminary facts about discrete, log-concave random variables.

For  $n = 1$ , the lower bound for  $H(X_1)$  given by Theorem 3 for the case of log-concave random variables on the integers is a significant improvement on the lower bound that can be obtained from the proof given in [17] for discrete random variables in a torsion-free group, which is  $\Omega\left(\frac{1}{\epsilon} \frac{1}{\epsilon}\right)$ .

Finally, let us note that Theorem 1 is a strong result: Although we suspect that the assumption of log-concavity may be relaxed, we do not expect it to hold in much greater generality; we believe that some structural conditions on the random variables should be necessary.

## 2. Notation and preliminaries

For a random variable  $X$  with p.m.f.  $p$  on the integers denote

$$q := \sum_{k \in \mathbb{Z}} \min\{p(k), p(k+1)\}. \tag{14}$$

The parameter  $q$  defined above plays an important role in a technique known as Bernoulli part decomposition, which has been used in [5, 12, 13] to prove local limit theorems. It was also used in [6] to prove the discrete entropic CLT mentioned in the Introduction.

In the present article, we use  $1 - q$  as a measure of smoothness of a p.m.f. on the integers. In what follows we will also write  $q(p)$  to emphasise the dependence on the p.m.f.  $p$ .

For two p.m.f.s on the integers  $p_1$  and  $p_2$ , we use the notation

$$\|p_1 - p_2\|_1 := \sum_{k \in \mathbb{Z}} |p_1(k) - p_2(k)|$$

for the  $\ell_1$ -distance between  $p_1$  and  $p_2$  and

$$\|p_1 - p_2\|_{TV} := \frac{1}{2} \|p_1 - p_2\|_1 \tag{15}$$

for the total variation distance.

**Proposition 4.** *Suppose  $X$  has p.m.f.  $p_X$  on  $\mathbb{Z}$  and let  $q = \sum_{k \in \mathbb{Z}} \min \{p_X(k), p_X(k + 1)\}$ . Then*

$$\|p_X - p_{X+1}\|_{TV} = 1 - q.$$

**Proof.** Since  $|a - b| = a + b - 2 \min \{a, b\}$ ,

$$\|p_X - p_{X+1}\|_1 = \sum_{k \in \mathbb{Z}} |p_X(k) - p_X(k + 1)| = 2 - 2 \sum_{k \in \mathbb{Z}} \min \{p_X(k), p_X(k + 1)\} = 2(1 - q).$$

The result follows. □

A p.m.f.  $p$  on  $\mathbb{Z}$  is called *log-concave*, if for any  $k \in \mathbb{Z}$

$$p(k)^2 \geq p(k - 1)p(k + 1). \tag{16}$$

If a random variable  $X$  is distributed according to a log-concave p.m.f. we say that  $X$  is log-concave. Throughout we suppose that  $X_1, \dots, X_n$  are i.i.d. random variables having a log-concave p.m.f. on the integers,  $p$ , common variance  $\sigma^2$  and denote their sum with  $S_n$ . Also, we denote

$$p_{\max} = p_{\max}(X) := \sup_k \mathbb{P}(X = k)$$

and write

$$N_{\max} = N_{\max}(X) := \max\{k \in \mathbb{Z} : p(k) = p_{\max}\}, \tag{17}$$

i.e.  $N_{\max}$  is the last  $k \in \mathbb{Z}$  for which the maximum probability is achieved. We will make use of the following bound from [3]:

**Lemma 5.** *Suppose  $X$  has discrete log-concave distribution with  $\sigma^2 = \text{Var}(X) \geq 1$ . Then*

$$\frac{1}{4\sigma} \leq p_{\max} \leq \frac{1}{\sigma}. \tag{18}$$

**Proof.** Follows immediately from [3, Theorem 1.1.]. □

**Proposition 6.** *Let  $X$  be a log-concave random variable on the integers with mean  $\mu \in \mathbb{R}$  and variance  $\sigma^2$ , and let  $\delta > 0$ . Then, if  $\sigma > 4^{1/2\delta}$ ,*

$$|N_{\max} - \mu| < \sigma^{3/2+\delta} + 1. \tag{19}$$

**Proof.** Suppose for contradiction that  $|N_{\max} - \mu| \geq \sigma^{3/2+\delta} + 1$ . Then, using (18),

$$\mathbb{P}(|X - \mu| > \sigma^{3/2+\delta}) \geq p(N_{\max}) \geq \frac{1}{4\sigma}.$$

But Chebyshev’s inequality implies

$$\mathbb{P}(|X - \mu| > \sigma^{3/2+\delta}) \leq \frac{1}{\sigma^{1+2\delta}} < \frac{1}{4\sigma}. \tag{20}$$

Below we show that for any integer-valued random variable  $X$ ,  $q \rightarrow 1$  implies  $H(X) \rightarrow \infty$ . It is not hard to see that the converse is not always true, i.e.  $H(X) \rightarrow \infty$  does not necessarily imply  $q \rightarrow 1$ : Consider a random variable with a mass of  $\frac{1}{2}$  at zero and all other probabilities equal on an increasingly large subset of  $\mathbb{Z}$ . Nevertheless, using Lemma 5, we show that if  $X$  is log-concave this implication is true. In fact, part 2 of Proposition 7 holds for all unimodal distributions. Clearly any

log-concave distribution is unimodal, since (16) is equivalent to the sequence  $\{\frac{p_X(k+1)}{p_X(k)}\}_{k \in \mathbb{Z}}$  being non-increasing.

**Proposition 7.** *Suppose that the random variable  $X$  has p.m.f.  $p_X$  on the integers and let  $q = q(p_X)$  as above. Then*

1.  $e^{-H(X)} \leq 1 - q$ .
2. *If  $p_X$  is unimodal, then  $1 - q = p_{\max}$ .*

**Proof.** Let  $m$  be a mode of  $X$ , that is  $p(m) = p_{\max}$ . Then

$$\begin{aligned} q &= \sum_{k \leq m-1} \min \{p_X(k), p_X(k+1)\} + \sum_{k \geq m} \min \{p_X(k), p_X(k+1)\} \\ &\leq \sum_{k \leq m-1} p_X(k) + \sum_{k \geq m} p_X(k+1) \\ &= 1 - p_{\max}. \end{aligned}$$

The bound 1 follows since  $H(X) = \mathbb{E}(\log \frac{1}{p_X(X)}) \geq \log \frac{1}{\max_k p_X(k)}$ .

For 2, note that since  $p_X$  is unimodal,  $p_X(k+1) \geq p_X(k)$  for all  $k < m$  and  $p_X(k+1) \leq p_X(k)$  for all  $k \geq m$ . Therefore, the inequality in part 1 is equality and 2 follows.  $\square$

### 3. Proofs of Theorems 1 and 3

Let  $U^{(n)} := \sum_{i=1}^n U_i$ , where  $U_i$  are i.i.d. continuous uniforms on  $(0, 1)$ . Let  $f_{S_n+U^{(n)}}$  denote the density of  $S_n + U^{(n)}$ . We approximate  $f_{S_n+U^{(n)}}$  with the p.m.f., say  $p_{S_n}$ , of  $S_n$ .

We recall that the class of discrete log-concave distributions is closed under convolution [10] and hence the following lemma may be applied to  $S_n$ .

**Lemma 8.** *Let  $S$  be a log-concave random variable on the integers with variance  $\sigma^2 = \text{Var}(S)$  and, for any  $n \geq 1$ , denote by  $f_{S+U^{(n)}}$  the density of  $S + U^{(n)}$  on the real line. Then for any  $n \geq 1$  and  $x \in \mathbb{R}$ ,*

$$f_{S+U^{(n)}}(x) = p_S(\lfloor x \rfloor) + g_n(\lfloor x \rfloor, x), \tag{20}$$

for some  $g_n : \mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{R}$  satisfying

$$\sum_{k \in \mathbb{Z}} \sup_{u \in [k, k+1)} |g_n(k, u)| \leq (2^n - 2) \frac{1}{\sigma}. \tag{21}$$

Moreover, if  $\lfloor x \rfloor \geq N_{\max} + n - 1$ ,

$$f_{S+U^{(n)}}(x) \leq 2^n p_S(\lfloor x \rfloor - n + 1). \tag{22}$$

**Proof.** First we recall that for a discrete random variable  $S$  and a continuous independent random variable  $U$  with density  $f_U$ ,  $S + U$  is continuous with density

$$f_{S+U}(x) = \sum_{k \in \mathbb{Z}} p_S(k) f_U(x - k).$$

For  $n = 1$ , the statement is true with  $g_n = 0$ . We proceed by induction on  $n$  with  $n = 2$  as base case, which illustrates the idea better. The density of  $U_1 + U_2$  is  $f_{U_1+U_2}(u) = u$ , for  $u \in (0, 1)$  and  $f_{U_1+U_2}(u) = 2 - u$ , for  $u \in [1, 2)$ . Thus, we have

$$f_{S+U^{(2)}}(x) = p_S(\lfloor x \rfloor) + (1 - x + \lfloor x \rfloor)(p_S(\lfloor x \rfloor - 1) - p_S(\lfloor x \rfloor)). \tag{23}$$

Therefore,

$$f_{S+U^{(2)}}(x) = p_S(\lfloor x \rfloor) + g_2(\lfloor x \rfloor, x),$$

where  $g_2(k, x) = (1 - x + k)(p_S(k - 1) - p_S(k))$  and by Propositions 4, 7.2 and Lemma 5

$$\sum_{k \in \mathbb{Z}} \sup_{u \in [k, k+1)} |g_2(k, u)| \leq \sum_{k \in \mathbb{Z}} |p_S(k) - p_S(k - 1)| = \|p_S - p_{S+1}\|_1 \leq \frac{2}{\sigma}.$$

Next, we have

$$\begin{aligned} f_{S+U^{(n+1)}}(x) &= \int_{(0,1)} f_{S+U^{(n)}}(x - u) du \\ &= \int_{(0,1) \cap (x - \lfloor x \rfloor - 1, x - \lfloor x \rfloor)} f_{S+U^{(n)}}(x - u) du + \int_{(0,1) \cap (x - \lfloor x \rfloor, x - \lfloor x \rfloor + 1)} f_{S+U^{(n)}}(x - u) du. \end{aligned} \tag{24}$$

Using the inductive hypothesis, (24) is equal to

$$\begin{aligned} &(x - \lfloor x \rfloor)p_S(\lfloor x \rfloor) + (1 - x + \lfloor x \rfloor)p_S(\lfloor x \rfloor - 1) \\ &+ \int_{(0,1) \cap (x - \lfloor x \rfloor - 1, x - \lfloor x \rfloor)} g_n(\lfloor x \rfloor, x - u) du + \int_{(0,1) \cap (x - \lfloor x \rfloor, x - \lfloor x \rfloor + 1)} g_n(\lfloor x \rfloor - 1, x - u) du \end{aligned} \tag{25}$$

with  $g_n$  satisfying (21). Thus, we can write

$$\begin{aligned} f_{S+U^{(n+1)}}(x) &= p_S(\lfloor x \rfloor) + (1 - x + \lfloor x \rfloor)(p_S(\lfloor x \rfloor - 1) - p_S(\lfloor x \rfloor)) \\ &+ \int_{(0,1) \cap (x - \lfloor x \rfloor - 1, x - \lfloor x \rfloor)} g_n(\lfloor x \rfloor, x - u) du + \int_{(0,1) \cap (x - \lfloor x \rfloor, x - \lfloor x \rfloor + 1)} g_n(\lfloor x \rfloor - 1, x - u) du \end{aligned} \tag{26}$$

$$= p_S(\lfloor x \rfloor) + g_{n+1}(\lfloor x \rfloor, x), \tag{27}$$

where  $g_{n+1}(k, x) = (1 - x + k)(p_S(k - 1) - p_S(k)) + \int_{(0,1) \cap (x - k - 1, x - k)} g_n(k, x - u) du + \int_{(0,1) \cap (x - k, x - k + 1)} g_n(k - 1, x - u) du$ . Therefore, since  $g_n$  satisfies (21),

$$\sum_{k \in \mathbb{Z}} \sup_{u \in [k, k+1)} |g_{n+1}(k, u)| \leq \frac{2}{\sigma} + 2 \sum_{k \in \mathbb{Z}} \sup_{u \in [k, k+1)} |g_n(k, u)| \leq \frac{2}{\sigma} + 2(2^n - 2) \frac{1}{\sigma} = (2^{n+1} - 2) \frac{1}{\sigma}, \tag{28}$$

completing the inductive step and thus the proof of (21).

Inequality (22) may be proved in a similar way by induction: For  $n = 2$ , by (23)

$$f_{S+U^{(2)}}(x) \leq 2p_S(\lfloor x \rfloor - 1), \tag{29}$$

since  $p_S(\lfloor x \rfloor) \leq p_S(\lfloor x \rfloor - 1)$  for  $\lfloor x \rfloor \geq N_{\max} + 1$ .

By (24) and the inductive hypothesis

$$\begin{aligned} f_{S+U^{(n+1)}}(x) &= \int_{(0,1) \cap (x - \lfloor x \rfloor - 1, x - \lfloor x \rfloor)} f_{S+U^{(n)}}(x - u) du + \int_{(0,1) \cap (x - \lfloor x \rfloor, x - \lfloor x \rfloor + 1)} f_{S+U^{(n)}}(x - u) du \end{aligned} \tag{30}$$

$$\leq 2^n p_S(\lfloor x \rfloor - n) + 2^n p_S(\lfloor x \rfloor - n + 1) \tag{31}$$

$$\leq 2^{n+1} p_S(\lfloor x \rfloor - n) \tag{32}$$

completing the proof of (22) and thus the proof of the lemma.  $\square$

**Lemma 9.** Let  $X$  be a log-concave random variable on the integers with p.m.f.  $p$ , mean zero and variance  $\sigma^2$ , and let  $0 < \epsilon < 1/2$ . If  $\sigma \geq \max\{3^{1/\epsilon}, (12e^3)^{1/(1-2\epsilon)}\}$ , there is an  $N_0 \in \{N_{\max}, \dots, N_{\max} + 2\lceil\sigma^2\rceil\}$  such that, for each  $k \geq N_0$ ,

$$p(k + 1) \leq \left(1 - \frac{1}{\sigma^{2-\epsilon}}\right)p(k). \tag{33}$$

Similarly, there is an  $N_0^- \in \{N_{\max} - 2\lceil\sigma^2\rceil, \dots, N_{\max}\}$  such that, for each  $k \leq N_0^-$ ,

$$p(k - 1) \leq \left(1 - \frac{1}{\sigma^{2-\epsilon}}\right)p(k).$$

**Proof.** Let  $\theta = 1 - \frac{1}{\sigma^{2-\epsilon}}$ . It suffices to show that there is an  $N_0 \in \{N_{\max}, \dots, N_{\max} + 2\lceil\sigma^2\rceil\}$  such that  $p(N_0 + 1) \leq \theta p(N_0)$ , since then, for each  $k \geq N_0$ ,  $\frac{p(k+1)}{p(k)} \leq \frac{p(N_0+1)}{p(N_0)} \leq \theta$  by log-concavity.

Suppose for contradiction that  $p(k + 1) \geq \theta p(k)$  for each  $k \in \{N_{\max}, \dots, N_{\max} + 2\lceil\sigma^2\rceil\}$ . Then, we have, using (18)

$$\sigma^2 = \sum_{k \in \mathbb{Z}} k^2 p(k) \geq \sum_{k=N_{\max}}^{N_{\max}+2\lceil\sigma^2\rceil} k^2 p(k) \geq \sum_{k=N_{\max}}^{N_{\max}+2\lceil\sigma^2\rceil} k^2 \theta^{k-N_{\max}} \frac{1}{4\sigma} \tag{34}$$

$$= \sum_{m=0}^{2\lceil\sigma^2\rceil} (N_{\max} + m)^2 \theta^m \frac{1}{4\sigma} \geq \sum_{m=\max\{0, -N_{\max}\}}^{2\lceil\sigma^2\rceil} (N_{\max} + m)^2 \theta^m \frac{1}{4\sigma}. \tag{35}$$

Now we use Proposition 6 with  $\delta > 0$  to be chosen later. Thus, the right-hand side of (35) is at least

$$\begin{aligned} & \sum_{m=\lceil\sigma^{3/2+\delta}+1\rceil}^{2\lceil\sigma^2\rceil} (N_{\max} + m)^2 \theta^m \frac{1}{4\sigma} \\ & \geq \sum_{m=\lceil\sigma^{3/2+\delta}+1\rceil}^{2\lceil\sigma^2\rceil} (m - \lceil\sigma^{3/2+\delta}\rceil - 1)^2 \theta^m \frac{1}{4\sigma} \end{aligned} \tag{36}$$

$$= \sum_{k=0}^{2\lceil\sigma^2\rceil - \lceil\sigma^{3/2+\delta}\rceil - 1} k^2 \theta^{k + \lceil\sigma^{3/2+\delta}\rceil + 1} \frac{1}{4\sigma} \tag{37}$$

$$\geq \theta^{\lceil\sigma^{3/2+\delta}\rceil + 1} \frac{1}{4\sigma} \sum_{k=1}^{\lceil\sigma^2\rceil - 1} k \theta^k \tag{38}$$

$$= \frac{\theta^{\lceil\sigma^{3/2+\delta}\rceil + 1}}{4\sigma} \left[ \theta \frac{1 - \theta^{\lceil\sigma^2\rceil}}{(1 - \theta)^2} - \lceil\sigma^2\rceil \frac{\theta^{\lceil\sigma^2\rceil}}{(1 - \theta)} \right]. \tag{39}$$

Using the elementary bound  $(1 - x)^y \geq e^{-2xy}$ , for  $0 < x < \frac{\log 2}{2}, y > 0$ , we see that

$$\theta^{\lceil\sigma^{3/2+\delta}\rceil + 1} = \left(1 - \frac{1}{\sigma^{2-\epsilon}}\right)^{\lceil\sigma^{3/2+\delta}\rceil + 1} \geq e^{-2\frac{\sigma^{3/2+\delta}+2}{\sigma^{2-\epsilon}}} \geq e^{-3},$$



where the last inequality holds as long as  $\epsilon + \delta < 1/2$ . Choosing  $\delta = 1/4 - \epsilon/2$ , we see that the assumption of Proposition 6 is satisfied for  $\sigma > 16^{1/(1-2\epsilon)}$  and thus for  $\sigma > (12e^3)^{1/(1-2\epsilon)}$  as well. Furthermore, using  $(1 - x)^y \leq e^{-xy}$ ,  $0 < x < 1, y > 0$ , we get  $\theta^{\sigma^2} \leq e^{-\sigma^\epsilon}$ . Thus, the right-hand side of (39) is at least

$$\begin{aligned} & \frac{1}{4e^3\sigma} \left[ \left(1 - \frac{1}{\sigma^{2-\epsilon}}\right) (1 - e^{-\sigma^\epsilon}) \sigma^{4-2\epsilon} - \sigma^{4-\epsilon} e^{-\sigma^\epsilon} - \sigma^{2-\epsilon} e^{-\sigma^\epsilon} \right] \\ & \geq \frac{\sigma^{3-2\epsilon}}{4e^3} \left[ 1 - 2 \frac{\sigma^\epsilon}{e^{\sigma^\epsilon}} - \frac{1}{\sigma^{2-\epsilon}} \right] \\ & > \sigma^2 \frac{\sigma^{1-2\epsilon}}{12e^3} \tag{40} \\ & \geq \sigma^2, \tag{41} \end{aligned}$$

where (40) holds for  $\sigma > 3^{\frac{1}{\epsilon}}$ , since then  $\frac{\sigma^\epsilon}{e^{\sigma^\epsilon}} \leq \frac{1}{4}$  and  $\sigma^{-2+\epsilon} < \sigma^{-1} < \frac{1}{9}$ . Finally, (41) holds for  $\sigma > (12e^3)^{\frac{1}{1-2\epsilon}}$ , getting the desired contradiction.

For the second part, apply the first part to the log-concave random variable  $-X$ . □

**Remark 10.** The bound (33) may be improved due to the suboptimal step (38), e.g. by means of the identity

$$\sum_{k=1}^M k^2 \theta^k = \theta \frac{d}{d\theta} \left( \sum_{k=1}^M k \theta^k \right) = \theta \frac{d}{d\theta} \left( \theta \frac{d}{d\theta} \left( \sum_{k=1}^M \theta^k \right) \right) = \theta \frac{d}{d\theta} \left( \theta \frac{d}{d\theta} \left( \frac{\theta - \theta^{M+1}}{1 - \theta} \right) \right).$$

It is, however, sufficient for our purpose as it will only affect a higher-order term in the proof of Theorem 1.

We are now ready to give the proof of Theorem 1 and of our main result, Theorem 3.

**Proof of Theorem 1.** Assume without loss of generality that  $X_1$  has zero mean. Let  $F(x) = x \log \frac{1}{x}, x > 0$  and note that  $F(x)$  is non-decreasing for  $x \leq 1/e$ . As before denote  $S_n = \sum_{i=1}^n X_i, U^{(n)} = \sum_{i=1}^n U_i$  and let  $f_{S_n+U^{(n)}}$  be the density of  $S_n + U^{(n)}$  on the reals. We have

$$\begin{aligned} & h(X_1 + \dots + X_n + U_1 + \dots + U_n) \\ & = \sum_{k \in (-5n\sigma^2, 5n\sigma^2)} \int_{[k, k+1)} F(f_{S_n+U^{(n)}}(x)) dx + \sum_{|k| \geq 5n\sigma^2} \int_{[k, k+1)} F(f_{S_n+U^{(n)}}(x)) dx. \tag{42} \end{aligned}$$

First, we will show that the “entropy tails”, i.e. the second term in (42), vanish as  $\sigma^2$  grows large. To this end, note that for  $k \geq 5n\sigma^2$ , we have  $p_{S_n}(k+1) \leq p_{S_n}(k)$ , since by Proposition 6 applied to the log-concave random variable  $S_n, N_{\max} \leq n\sigma^2 + 1$  as long as  $\sqrt{n}\sigma > 4$ . Thus, by (22), for  $k \geq 5n\sigma^2$  and  $x \in [k, k+1), f_{S_n+U^{(n)}}(x) \leq 2^n p_{S_n}(k-n+1)$ . Hence, for

$$\sigma > \frac{2^n}{\sqrt{n}} e, \tag{43}$$

we have, using the monotonicity of  $F$  for  $x \leq \frac{1}{e}$ ,

$$0 \leq \sum_{k \geq 5n\sigma^2} \int_{[k, k+1]} F(f_{S_n+U^{(n)}}(x)) dx \leq \sum_{k \geq 5n\sigma^2} F(2^n p_{S_n}(k-n+1)) \tag{44}$$

$$= \sum_{k \geq 5n\sigma^2} 2^n p_{S_n}(k-n+1) \log \frac{1}{2^n p_{S_n}(k-n+1)} \tag{45}$$

$$\leq 2^n \frac{1}{\sqrt{n\sigma}} \sum_{k \geq 5n\sigma^2} \theta^{k-4\lceil n\sigma^2 \rceil} \log \frac{\sqrt{n\sigma}}{2^n \theta^{k-4\lceil n\sigma^2 \rceil}} \tag{46}$$

$$= 2^n \frac{\log \frac{1}{\theta}}{\sqrt{n\sigma}} \sum_{m \geq n\sigma^2} m \theta^m + 2^n \frac{\log \frac{\sqrt{n\sigma}}{2^n}}{\sqrt{n\sigma}} \sum_{m \geq n\sigma^2} \theta^m \tag{47}$$

$$\leq 2^{n+1} \frac{\log \sqrt{n\sigma}}{\sqrt{n\sigma}} \left[ \frac{\theta^{\lceil n\sigma^2 \rceil + 1}}{(1-\theta)^2} + \lceil n\sigma^2 \rceil \frac{\theta^{\lceil n\sigma^2 \rceil}}{1-\theta} + \frac{\theta^{\lceil n\sigma^2 \rceil}}{1-\theta} \right] \tag{48}$$

$$\leq 2^{n+1} \frac{\log \sqrt{n\sigma}}{\sqrt{n\sigma}} e^{-(\sqrt{n\sigma})^\epsilon} \left[ (\sqrt{n\sigma})^{4-2\epsilon} + (n\sigma^2 + 1)(\sqrt{n\sigma})^{2-\epsilon} + (\sqrt{n\sigma})^{2-\epsilon} \right] \tag{49}$$

$$\leq 2^{n+3} \frac{\log \sqrt{n\sigma}}{\sqrt{n\sigma}} e^{-(\sqrt{n\sigma})^\epsilon} (\sqrt{n\sigma})^{4-\epsilon} \tag{50}$$

$$\leq 2^{n+4} e^{-(\sqrt{n\sigma})^{1/5}} (\sqrt{n\sigma})^3. \tag{51}$$

Here, (46) holds for

$$\sqrt{n\sigma} > 3^7 \tag{52}$$

with  $\theta = 1 - \frac{1}{(\sqrt{n\sigma})^{2-\epsilon}} = 1 - \frac{1}{(\sqrt{n\sigma})^{9/5}}$ , where we have used Lemma 9 with  $\epsilon = 1/5$  (which makes the assumption approximately minimal). In particular, repeated application of (33) yields

$$p_{S_n}(k-n+1) \leq \theta^{k-4\lceil n\sigma^2 \rceil} p_{S_n}(4\lceil n\sigma^2 \rceil - n + 1) \leq \frac{\theta^{k-4\lceil n\sigma^2 \rceil}}{\sqrt{n\sigma}}.$$

We bound the left tail in the exact the same way, using the second part of Lemma 9:

$$0 \leq \sum_{k \leq -5n\sigma^2} \int_{[k, k+1]} F(f_{S_n+U^{(n)}}(x)) dx \leq 2^{n+4} e^{-(\sqrt{n\sigma})^{1/5}} (\sqrt{n\sigma})^3. \tag{53}$$

Next we will show that the first term in (42) is approximately  $H(S_n)$  to complete the proof. We have

$$\begin{aligned} & \sum_{k \in (-5n\sigma^2, 5n\sigma^2)} \int_{[k, k+1]} F(f_{S_n+U^{(n)}}(x)) dx = \log n\sigma^2 \int_{(-\lfloor 5n\sigma^2 \rfloor, \lfloor 5n\sigma^2 \rfloor + 1)} f_{S_n+U^{(n)}}(x) dx \\ & + \sum_{k \in (-5n\sigma^2, 5n\sigma^2)} \int_{[k, k+1]} F(f_{S_n+U^{(n)}}(x)) - f_{S_n+U^{(n)}}(x) \log n\sigma^2 dx \end{aligned} \tag{54}$$

$$\begin{aligned} & = \log n\sigma^2 \mathbb{P}(S_n + U^{(n)} \in (-\lfloor 5n\sigma^2 \rfloor, \lfloor 5n\sigma^2 \rfloor + 1)) \\ & + \sum_{k \in (-5n\sigma^2, 5n\sigma^2)} \int_{[k, k+1]} F(f_{S_n+U^{(n)}}(x)) - f_{S_n+U^{(n)}}(x) \log n\sigma^2 dx. \end{aligned} \tag{55}$$

Now we will apply the estimate of Lemma 12, which is stated and proved in the Appendix, to the integrand of the second term in (54) with  $G(x) = F(x) - x \log(n\sigma^2)$ ,  $\mu = \frac{1}{11\sigma\sqrt{n}}$ ,  $D =$

$2^n \sqrt{n}\sigma$ ,  $M = n\sigma^2$ ,  $a = f_{S_n+U^{(n)}}(x)$  and  $b = p_{S_n}(k)$ . We obtain, using Lemma 8,

$$\begin{aligned} & \sum_{k \in (-5n\sigma^2, 5n\sigma^2)} \left| \int_{[k, k+1)} G(f_{S_n+U^{(n)}}(x)) dx - G(p_{S_n}(k)) \right| \\ & \leq 11n\sigma^2 \frac{2 \log(11\sigma \sqrt{n})}{11\sigma^3 n \sqrt{n}} + \sum_{k \in \mathbb{Z}} \int_{[k, k+1)} |f_{S_n+U^{(n)}}(x) - p_{S_n}(k)| dx \left( \log(11\sigma \sqrt{n}) + \log(e2^n \sigma \sqrt{n}) \right) \end{aligned} \tag{56}$$

$$\leq \frac{2 \log(11\sigma \sqrt{n})}{\sigma \sqrt{n}} + \sum_{k \in \mathbb{Z}} \sup_{x \in [k, k+1)} |g_n(k, x)| \left( \log(11\sigma \sqrt{n}) + \log(e2^n \sigma \sqrt{n}) \right) \tag{57}$$

$$\leq \frac{2 \log(11\sigma \sqrt{n})}{\sigma \sqrt{n}} + \frac{2^{n+1}}{\sigma \sqrt{n}} \log(11e2^n \sigma \sqrt{n}) \tag{58}$$

$$\leq \frac{3}{4} \frac{2^{n+2}}{\sigma \sqrt{n}} \log(11e2^n \sigma \sqrt{n}) \tag{59}$$

$$\leq \frac{2^{n+2}}{\sigma \sqrt{n}} \log((11e)^{3/4} 2^n \sigma^{3/4} \sqrt{n}) \leq \frac{2^{n+2}}{\sigma \sqrt{n}} \log(2^{n+2} \sigma \sqrt{n}), \tag{60}$$

where  $g_n(k, x)$  is given by Lemma 8 applied to the log-concave random variable  $S_n$  and therefore  $\sum_k \sup_{x \in [k, k+1)} g_n(k, x) \leq \frac{2^n}{\sigma \sqrt{n}}$ . In the last inequality in (60) we have used that  $\frac{(11e)^{3/4}}{\sigma^{1/4}} \leq 4$ , for  $\sigma > 3^7$ . Therefore, by (54) and (60),

$$\begin{aligned} & \left| H(S_n) - \sum_{k \in (-5n\sigma^2, 5n\sigma^2)} \int_{[k, k+1)} F(f_{S_n+U^{(n)}}(x)) dx \right| \\ & \leq \left| H(S_n) - \sum_{k \in (-5n\sigma^2, 5n\sigma^2)} G(p_{S_n}(k)) - \log n\sigma^2 \mathbb{P}(S_n + U^{(n)} \in (-\lfloor 5n\sigma^2 \rfloor, \lfloor 5n\sigma^2 \rfloor + 1)) \right| \\ & \quad + \frac{2^{n+2}}{\sigma \sqrt{n}} \log(2^{n+2} \sigma \sqrt{n}) \end{aligned} \tag{61}$$

$$\begin{aligned} & \leq \left| H(S_n) - \sum_{k \in (-5n\sigma^2, 5n\sigma^2)} F(p_{S_n}(k)) \right| + \frac{2^{n+2}}{\sigma \sqrt{n}} \log(2^{n+2} \sigma \sqrt{n}) \\ & \quad + \log n\sigma^2 \left| \mathbb{P}(S_n + U^{(n)} \in (-\lfloor 5n\sigma^2 \rfloor, \lfloor 5n\sigma^2 \rfloor + 1)) - \mathbb{P}(S_n \in (-5n\sigma^2, 5n\sigma^2 + 1)) \right| \end{aligned} \tag{62}$$

$$\begin{aligned} & \leq \sum_{|k| \geq 5n\sigma^2} F(p_{S_n}(k)) + \frac{2^{n+2}}{\sigma \sqrt{n}} \log(2^{n+2} \sigma \sqrt{n}) \\ & \quad + \log n\sigma^2 \left| \mathbb{P}(S_n + U^{(n)} \in (-\lfloor 5n\sigma^2 \rfloor, \lfloor 5n\sigma^2 \rfloor + 1)) - \mathbb{P}(S_n \in (-5n\sigma^2, 5n\sigma^2 + 1)) \right|. \end{aligned} \tag{63}$$

But, in view of (45), we can bound the discrete tails in the same way:

$$\sum_{|k| \geq 5n\sigma^2} F(p_{S_n}(k)) \leq 2^{n+5} e^{-(\sqrt{n}\sigma)^{1/5}} (\sqrt{n}\sigma)^3. \tag{64}$$

Finally, note that by Chebyshev’s inequality

$$0 \leq \mathbb{P}(S_n + U^{(n)} \notin (-\lfloor 5n\sigma^2 \rfloor, \lfloor 5n\sigma^2 \rfloor + 1)) \leq \mathbb{P}\left(|S_n + U^{(n)} - \frac{n}{2}| > 4n\sigma^2\right) \leq \frac{1}{8n\sigma^2} \tag{65}$$

and the same upper bound applies to  $\mathbb{P}(S_n \notin (-5n\sigma^2, 5n\sigma^2 + 1))$ . Since both probabilities inside the absolute value in (63) are also upper bounded by 1, replacing the bounds (65) and (66) into (63), we get

$$\begin{aligned} & \left| H(S_n) - \sum_{k \in (-5n\sigma^2, 5n\sigma^2)} \int_{[k, k+1)} F(f_{S_n+U^{(n)}}(x)) dx \right| \\ & \leq 2^{n+5} e^{-(\sqrt{n}\sigma)^{1/5}} (\sqrt{n}\sigma)^3 + \frac{2^{n+2}}{\sigma\sqrt{n}} \log(2^{n+2}\sigma\sqrt{n}) + \frac{\log n\sigma^2}{8n\sigma^2}. \end{aligned} \tag{66}$$

In view of (42) and the bounds on the continuous tails (51), (53), we conclude

$$|h(S_n + U^{(n)}) - H(S_n)| \leq 2^{n+6} e^{-(\sqrt{n}\sigma)^{1/5}} (\sqrt{n}\sigma)^3 + \frac{2^{n+2}}{\sigma\sqrt{n}} \log(2^{n+2}\sigma\sqrt{n}) + \frac{\log n\sigma^2}{8n\sigma^2} \tag{67}$$

as long as (43) and (52) are satisfied, that is as long as  $\sigma > \max\{2^{n+2}/\sqrt{n}, 3^7/\sqrt{n}\}$ . □

**Remark 11.** The exponent in (51) can be improved due to the suboptimal step (38) in Lemma 9. However, this is only a third-order term and therefore the rate in Theorem 1 would still be of the same order.

**Proof of Theorem 3.** Let  $U_1, \dots, U_n$  be continuous i.i.d. uniforms on  $(0, 1)$ . Then by the generalised EPI for continuous random variables [1, 11]

$$h\left(\frac{X_1 + \dots + X_{n+1} + U_1 + \dots + U_{n+1}}{\sqrt{n+1}}\right) \geq h\left(\frac{X_1 + \dots + X_n + U_1 + \dots + U_n}{\sqrt{n}}\right). \tag{68}$$

But by the scaling property of differential entropy [4], this is equivalent to

$$h(X_1 + \dots + X_{n+1} + U_1 + \dots + U_{n+1}) \geq h(X_1 + \dots + X_n + U_1 + \dots + U_n) + \frac{1}{2} \log\left(\frac{n+1}{n}\right). \tag{69}$$

Now we claim that for every  $n \geq 1$ , if  $H(X_1) \geq \log \frac{2}{\epsilon} + \log \log \frac{2}{\epsilon} + n + 26$  then

$$|h(X_1 + \dots + X_n + U_1 + \dots + U_n) - H(X_1 + \dots + X_n)| \leq \frac{\epsilon}{2}. \tag{70}$$

Then the result follows from (71), applied to both sides of (70) (for  $n$  and  $n + 1$ , respectively).

To prove the claim (71), we invoke Theorem 1. To this end let  $n \geq 1$  and assume that  $H(X_1) \geq \log \frac{2}{\epsilon} + \log \log \frac{2}{\epsilon} + n + 26$ . First we note, that since [4]

$$H(X_1) = h(X_1 + U_1) \leq \frac{1}{2} \log(2\pi e(\sigma^2 + 1/12)), \tag{71}$$

we have  $e^{H(X_1)} \leq 6\sigma$  provided that  $\sigma > 0.275$ . Thus,  $H(X_1) \geq 26$  implies  $\sigma > 50^6 > 90^5$ . Therefore the assumptions of the theorem are satisfied and we get

$$\begin{aligned} & |h(X_1 + \dots + X_n + U_1 + \dots + U_n) - H(X_1 + \dots + X_n)| \\ & \leq 2^{n+6} e^{-(\sqrt{n}\sigma)^{1/5}} (\sqrt{n}\sigma)^3 + \frac{2^{n+2}}{\sigma\sqrt{n}} \log(2^{n+2}\sigma\sqrt{n}) + \frac{\log n\sigma^2}{8n\sigma^2} \end{aligned} \tag{72}$$

$$\leq (2^6 + 2^3 + 1) 2^n \frac{\log \sqrt{n}\sigma}{\sqrt{n}\sigma} = 73 \cdot 2^n \frac{\log \sqrt{n}\sigma}{\sqrt{n}\sigma}. \tag{73}$$

In (74), we used the elementary fact that for  $x \geq 90^5$ ,  $\frac{x^3}{e^{x^{1/5}}} \leq \frac{1}{x} \leq \frac{\log x}{x}$  to bound the first term and the assumption  $\sigma\sqrt{n} \geq 2^{n+2}$  to bound the second term. Thus, by assumption  $\sigma \geq \frac{e^{H(X_1)}}{6} \geq$

$\frac{2}{\epsilon} \log \frac{2}{\epsilon} e^{n+24}$  and since  $\frac{\log x}{x}$  is non-increasing for  $x > e$ , we obtain by (74)

$$|h(X_1 + \dots + X_n + U_1 + \dots + U_n) - H(X_1 + \dots + X_n)| \tag{74}$$

$$\leq 73 \cdot 2^n \frac{\log \frac{2}{\epsilon} + \log \log \frac{2}{\epsilon} + n + 24}{\frac{2}{\epsilon} \log \frac{2}{\epsilon} e^{n+24}} \tag{75}$$

$$\leq \frac{\epsilon}{2} \left[ \frac{146}{(\frac{\epsilon}{2})^n e^{24}} + \frac{n73}{\log \frac{2}{\epsilon} (\frac{\epsilon}{2})^n e^{24}} + \frac{1752}{\log \frac{2}{\epsilon} (\frac{\epsilon}{2})^n e^{24}} \right] \tag{76}$$

$$< \frac{\epsilon}{2} \tag{77}$$

proving the claim (71) and thus the theorem. □

### Acknowledgements

The author is indebted to Ioannis Kontoyiannis for interesting discussions as well as many useful suggestions and comments. The author would also like to thank the two anonymous reviewers for the careful reading of the manuscript and for many useful comments that significantly improved the presentation of the results in the paper.

### References

- [1] Artstein, S., Ball, K., Barthe, F. and Naor, A. (2004) Solution of Shannon’s problem on the monotonicity of entropy. *J. Am. Math. Soc.* **17**(4) 975–982.
- [2] Barron, A. R. (1986) Entropy and the central limit theorem, *Ann. Prob.* **14**, 336–342.
- [3] Bobkov, S. G., Marsiglietti, A. and Melbourne, J. (2022) Concentration functions and entropy bounds for discrete log-concave distributions. *Comb. Prob. Comput.* **31**(1) 54–72.
- [4] Cover, T. M. and Thomas, J. A. (2006) *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA, *Wiley Series in Telecommunications and Signal Processing*.
- [5] Davis, B. and McDonald, D. (1995) An elementary proof of the local central limit theorem. *J. Theor. Probab.* **8**(3) 693–701.
- [6] Gavalakis, L. and Kontoyiannis, I. (2021) Entropy and the discrete central limit theorem, arXiv preprint arXiv: 2106.00514.
- [7] Haghghatshoar, S., Abbe, E. and Telatar, E. (2012) Adaptive sensing using deterministic partial Hadamard matrices. IEEE, pp. 1842–1846. In *2012 IEEE International Symposium on Information Theory Proceedings*.
- [8] Haghghatshoar, S., Abbe, E. and Telatar, I. E. (2014) A new entropy power inequality for integer-valued random variables. *IEEE Trans. Inform. Theory* **60**(7) 3787–3796.
- [9] Harremoés, P. and Vignat, C. (2003) An entropy power inequality for the binomial family. *JIPAM. J. Inequal. Pure Appl. Math.* **4**(5) 93.
- [10] Hoggar, S. (1974) Chromatic polynomials and logarithmic concavity. *J. Comb. Theory, Ser. B* **16**(3) 248–254.
- [11] Madiman, M. and Barron, A. (2007) Generalized entropy power inequalities and monotonicity properties of information. *IEEE Trans. Inform. Theory* **53**(7) 2317–2329.
- [12] McDonald, D. R. (1980) On local limit theorem for integer-valued random variables. *Theory Prob. Appl.* **24**(3) 613–619.
- [13] Mineka, J. (1973) A criterion for tail events for sums of independent random variables. *Z. für Wahrscheinlichkeitstheorie und Verw. Gebiete* **25**(3) 163–170.
- [14] Ruzsa, I. Z. (2009) Sumsets and entropy. *Random Struct. Algorithms* **34**(1) 1–10.
- [15] Shannon, C. E. (1948) A mathematical theory of communication. *Bell Syst. Tech. J.* **27**(3) 379–423.
- [16] Stam, A. J. (1959) Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Control* **2**(2) 101–112.
- [17] Tao, T. (2010) Sumset and inverse sumset theory for Shannon entropy. *Comb. Probab. Comput.* **19**(4) 603–639.
- [18] Tao, T. and Vu, V. H. (2005) Entropy methods [Online]. Available: <http://www.math.ucla.edu/~tao/preprints/Expository/>.
- [19] Tao, T. and Vu, V. H. (2006) *Additive Combinatorics*. Cambridge University Press, *Cambridge Studies in Advanced Mathematics*.
- [20] Woo, J. O. and Madiman, M. (2015) A discrete entropy power inequality for uniform distributions. IEEE, pp. 1625–1629. In *2015 IEEE International Symposium on Information Theory (ISIT)*.

**Appendix**

**An elementary Lemma**

Here, we prove the following Taylor-type estimate that we used in the proof of Theorem 1. A similar estimate was used in [17].

**Lemma 12.** *Let  $D, M \geq 1$  and, for  $x > 0$ , consider  $G(x) = F(x) - x \log M$ , where  $F(x) = -x \log x$ . Then, for  $0 \leq a, b \leq \frac{D}{M}$  and any  $0 < \mu < \frac{1}{e}$ , we have the estimate*

$$|G(b) - G(a)| \leq \frac{2\mu}{M} \log \frac{1}{\mu} + |b - a| \left[ \log \frac{1}{\mu} + \log(eD) \right]. \tag{A.1}$$

**Proof.** Note that  $G'(x) = -\log x - 1 - \log M$ , which is non-negative for  $x < \frac{1}{eM}$ .

We will consider two cases separately.

The first case is when either  $a < \frac{\mu}{M}$  or  $b < \frac{\mu}{M}$ . Assume without loss of generality that  $a < \frac{\mu}{M}$ . Then if  $b < \frac{\mu}{M}$  as well, we have  $|G(b) - G(a)| \leq G(a) + G(b) \leq \frac{2\mu}{M} \log \frac{1}{\mu}$ , since then  $G' \geq 0$ . On the other hand, if  $b \geq \frac{\mu}{M} > a$  then  $G(a) \geq a \log \frac{1}{\mu}$  and  $G(b) \leq b \log \frac{1}{\mu}$ .

But then, either  $G(b) > G(a)$ , whence  $|G(b) - G(a)| \leq |b - a| \log \frac{1}{\mu}$  or  $G(b) < G(a)$ , whence  $|G(b) - G(a)| \leq |b - a| \log(eD)$ , since we must have  $G(a) - G(b) = (a - b)G'(\xi)$ , for some  $\xi \in (\frac{1}{eM}, \frac{D}{M}]$ .

Thus, in the first case,  $|G(b) - G(a)| \leq \frac{2\mu}{M} \log \frac{1}{\mu} + |b - a| \left[ \log \frac{1}{\mu} + \log(eD) \right]$ .

The second case, is when both  $a, b \geq \frac{\mu}{M}$ . Then  $G(b) - G(a) = (b - a)G'(\xi)$ , for some  $\frac{D}{M} \geq \xi \geq \mu \frac{1}{M}$ .

Since then  $|G'(\xi)| \leq \log \frac{1}{\mu} + \log D + 1$ , we have  $|G(b) - G(a)| \leq |b - a| \left( \log \frac{1}{\mu} + \log(eD) \right)$ .

In any case,  $|G(b) - G(a)| \leq \frac{2\mu}{M} \log \frac{1}{\mu} + |b - a| \left[ \log \frac{1}{\mu} + \log(eD) \right]$ . □