A reverse entropy power inequality for log-concave random vectors

Keith Ball, Piotr Nayar* and Tomasz Tkocz

8/06/2016

Abstract

We prove that the exponent of the entropy of one dimensional projections of a log-concave random vector defines a 1/5-seminorm. We make two conjectures concerning reverse entropy power inequalities in the log-concave setting and discuss some examples.

1   Introduction

One of the most significant and mathematically intriguing quantities studied in information theory is the *entropy*. For a random variable X with density f its entropy is defined as

$$\mathcal{S}(X) = \mathcal{S}(f) = -\int_{\mathbb{R}} f \ln f \tag{1}$$

provided this integral exists (in the Lebesgue sense). Note that the entropy is translation invariant and $\mathcal{S}(bX) = \mathcal{S}(X) + \ln|b|$ for any nonzero b. If f belongs to $L_p(\mathbb{R})$ for some $p > 1$, then by the concavity of the logarithm and Jensen's inequality $\mathcal{S}(f) > -\infty$. If $\mathbb{E}X^2 < \infty$, then comparison with the standard Gaussian density and again Jensen's inequality yields $\mathcal{S}(X) < \infty$. Particularly, the entropy of a log-concave random variable is well defined and finite. Recall that a random vector in $\mathbb{R}^n$ is called log-concave if it has a density of the form $e^{-\psi}$ with $\psi : \mathbb{R}^n \to (-\infty, +\infty]$ being a convex function.

The entropy power inequality (EPI) says that

$$e^{\frac{2}{n}\mathcal{S}(X+Y)} \geq e^{\frac{2}{n}\mathcal{S}(X)} + e^{\frac{2}{n}\mathcal{S}(Y)}, \tag{2}$$

for independent random vectors X and Y in $\mathbb{R}^n$ provided that all the entropies exist. Stated first by Shannon in his seminal paper [24] and first rigorously proved by Stam in [25] (see also [6]), it is often referred to as the Shannon-Stam inequality and plays a crucial role in information theory and elsewhere (see the survey [18]). Using the AM-GM inequality, the EPI can be *linearised*: for every $\lambda \in [0,1]$ and independent random vectors X, Y we have

$$\mathcal{S}(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \geq \lambda\mathcal{S}(X) + (1-\lambda)\mathcal{S}(Y) \tag{3}$$

provided that all the entropies exist. This formulation is in fact equivalent to (2) as first observed by Lieb in [22], where he also shows how to derive (3) from Young's inequality with sharp constants. Several other proofs of (3) are available, including versions for the Fisher information [13] and recent techniques of the minimum mean-square error [27]. There are also refinements when one variable is Gaussian [15], [17], [28].

If X and Y are independent and identically distributed random variables (or vectors), inequality (3) says that the entropy of the normalised sum

$$X_\lambda = \sqrt{\lambda}X + \sqrt{1-\lambda}Y \tag{4}$$

is at least as big as the entropy of the summands X and Y, $\mathcal{S}(X_\lambda) \geq \mathcal{S}(X)$. It is worth mentioning that this phenomenon has been quantified, first in [14], which has deep consequences in probability (see the pioneering work [4] and its sequels [1, 2] which establish the rate of convergence in the entropic central limit theorem and the "second law of probability" of the entropy growth, as well as the independent work [20], with somewhat different methods). In the context of log-concave vectors, Ball and Nguyen in [5] establish dimension free lower bounds on $\mathcal{S}(X_{1/2}) - \mathcal{S}(X)$ and discuss connections between the entropy and major conjectures in convex geometry; for the latter see also [12].

In general, the EPI cannot be reversed. In [7], Proposition V.8, Bobkov and Chistyakov find a random vector X with a finite entropy such that $\mathcal{S}(X+Y) = \infty$ for every independent of X random vector Y with finite entropy. However, for log-concave vectors and, more generally, convex measures, Bobkov and Madiman have recently addressed the question of reversing the EPI (see [10, 11]). They show that for any pair X, Y of independent log-concave random vectors in $\mathbb{R}^n$, there are linear volume preserving maps $T_1, T_2 : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$e^{\frac{2}{n}\mathcal{S}(T_1(X)+T_2(Y))} \leq C(e^{\frac{2}{n}\mathcal{S}(X)} + e^{\frac{2}{n}\mathcal{S}(Y)}),$$

where C is some universal constant.

For a random variable X with finite variance its *relative entropy* $\mathcal{D}(X)$ is defined as the difference $\mathcal{S}(Z) - \mathcal{S}(X)$, where Z is a Gaussian random variable with variance $\text{Var}(X)$. This quantity is nonnegative and provides a way to measure closeness to Gaussians. Another reverse EPI has been lately discovered in the context of the stability of Cramér's theorem (see [9]). The authors bound from below the relative entropy of the sum of independent *regularised* random variables in terms of the relative entropies of the regularised summands. The regularisation is performed by adding independent Gaussians and without it, such a lower bound does not hold in general, as shown by the same authors in [8].

The goal of this note is to investigate further in the log-concave setting some new forms of what could be called a reverse EPI. In the next section we present our results. The last section is devoted to their proofs.

Acknowledgements

2   Main results and conjectures

Suppose X is a symmetric log-concave random vector in $\mathbb{R}^n$. Then any projection of X on a certain direction $v \in \mathbb{R}^n$, that is the random variable $\langle X, v \rangle$ is also log-concave. Here $\langle \cdot, \cdot \rangle$ denotes the standard scalar product in $\mathbb{R}^n$. If we know the entropies of projections in, say two different directions, can we say anything about the entropy of projections in related directions? We make the following conjecture.

Conjecture 1. Let X be a symmetric log-concave random vector in $\mathbb{R}^n$. Then the function

$$N_X(v) = \begin{cases} e^{\mathcal{S}(\langle v, X \rangle)} & v \neq 0, \\ 0 & v = 0 \end{cases}$$

defines a norm on $\mathbb{R}^n$.

The homogeneity of N_X is clear. To check the triangle inequality, we have to answer really a two-dimensional question: *is it true that for a symmetric log-concave random vector (X, Y) in $\mathbb{R}^2$ we have*

$$e^{\mathcal{S}(X+Y)} \leq e^{\mathcal{S}(X)} + e^{\mathcal{S}(Y)}? \tag{5}$$

3

Indeed, this applied to the vector $(\langle u, X \rangle, \langle v, X \rangle)$ which is also log-concave yields $N_X(u+v) \leq N_X(u) + N_X(v)$. Inequality (5) can be seen as a reverse EPI, cf. (2). It is not too difficult to show that this inequality holds up to a multiplicative constant.

Proposition 1. *Let* (X, Y) *be a symmetric log-concave random vector on* $\mathbb{R}^2$. *Then*

$$e^{\mathcal{S}(X+Y)} \leq e \left(e^{\mathcal{S}(X)} + e^{\mathcal{S}(Y)} \right).$$

Proof. The argument relies on the well-known observation that for a log-concave density $f \colon \mathbb{R} \longrightarrow [0, +\infty)$ its maximum and entropy are related (see for example [5] or [12]),

$$-\ln \|f\|_\infty \leq \mathcal{S}(f) \leq 1 - \ln \|f\|_\infty. \tag{6}$$

Suppose that w is an even log-concave density of (X, Y). The densities of X, Y and $X + Y$ equal respectively

$$f(x) = \int w(x, t)\mathrm{d}t, \qquad g(x) = \int w(t, x)\mathrm{d}t, \qquad h(x) = \int w(x - t, t)\mathrm{d}t. \tag{7}$$

They are even and log-concave, hence attain their maximum at zero. By the result of Ball (Busemann's theorem for symmetric log-concave measures, see [3]), the function $\|x\|_w = (\int w(tx)\mathrm{d}t)^{-1}$ is a norm on $\mathbb{R}^2$. Particularly,

$$\frac{1}{\|h\|_\infty} = \frac{1}{h(0)} = \frac{1}{\int w(-t, t)\mathrm{d}t} = \|e_2 - e_1\|_w \leq \|e_1\|_w + \|e_2\|_w$$

$$= \frac{1}{\int w(t, 0)\mathrm{d}t} + \frac{1}{\int w(0, t)\mathrm{d}t} = \frac{1}{f(0)} + \frac{1}{g(0)} = \frac{1}{\|f\|_\infty} + \frac{1}{\|g\|_\infty}.$$

Using (6) twice we obtain

$$e^{\mathcal{S}(X+Y)} \leq \frac{e}{\|h\|_\infty} \leq e \cdot \left(\frac{1}{\|f\|_\infty} + \frac{1}{\|g\|_\infty} \right) \leq e \cdot \left(e^{\mathcal{S}(X)} + e^{\mathcal{S}(Y)} \right).$$

$\square$

Recall that the classical result of Aoki and Rolewicz says that a C-quasi-norm (1-homogeneous function satisfying the triangle inequality up to a multiplicative constant C) is equivalent to some κ-semi-norm (κ-homogeneous function satisfying the triangle inequality) for some κ depending only on C (to be precise, it is enough to take $\kappa = \ln 2 / \ln(2C)$). See for instance Lemma 1.1 and Theorem 1.2 in [21]. In view of Proposition 1, for every symmetric log-concave random vector X in $\mathbb{R}^n$ the function $N_X(v)^\kappa = e^{\kappa \mathcal{S}(\langle X, v \rangle)}$ with $\kappa = \frac{\ln 2}{1 + \ln 2}$ is equivalent to some nonnegative κ-semi-norm. Therefore, it is natural to relax Conjecture 1 and ask whether there is a positive universal constant κ such that the function N_X^κ itself satisfies the triangle inequality for every symmetric log-concave random vector X in $\mathbb{R}^n$. Our main result answers this question positively.

Theorem 1. *There exists a universal constant $\kappa > 0$ such that for a symmetric log-concave random vector X in $\mathbb{R}^n$ and two vectors $u, v \in \mathbb{R}^n$ we have*

$$e^{\kappa \mathcal{S}(\langle u+v, X \rangle)} \le e^{\kappa \mathcal{S}(\langle u, X \rangle)} + e^{\kappa \mathcal{S}(\langle v, X \rangle)}. \tag{8}$$

Equivalently, for a symmetric log-concave random vector (X, Y) in $\mathbb{R}^2$ we have

$$e^{\kappa \mathcal{S}(X+Y))} \le e^{\kappa \mathcal{S}(X)} + e^{\kappa \mathcal{S}(Y)}. \tag{9}$$

In fact, we can take $\kappa = 1/5$.

Remark 1. If we take X and Y to be independent random variables uniformly distributed on the intervals $[-t/2, t/2]$ and $[-1/2, 1/2]$ with $t < 1$, then (9) becomes $e^{\kappa t/2} \le 1 + t^\kappa$. Letting $t \to 0$ shows that necessarily $\kappa \le 1$. We believe that this is the extreme case and the optimal value of κ equals 1.

Remark 2. Inequality (9) with $\kappa = 1$ can be easily shown for log-concave random vectors (X, Y) in $\mathbb{R}^2$ for which one marginal has the same law as the other one rescaled, say $Y \sim tX$ for some $t > 0$. Note that the symmetry of (X, Y) is not needed here. This fact in the essential case of $t = 1$ was first observed in [16]. We recall the argument in the next section. Moreover, in that paper the converse was shown as well: given a density f, the equality

$$\max\{\mathcal{S}(X + Y), \ X \sim f, Y \sim f\} = \mathcal{S}(2X)$$

holds if and only if f is log-concave, thus characterizing log-concavity. For some bounds on $\mathcal{S}(X \pm Y)$ in higher dimensions see [23] and [11].

It will be much more convenient to prove Theorem 1 in an equivalent form, obtained by linearising inequality (9).

Theorem 2. *Let (X, Y) be a symmetric log-concave vector in $\mathbb{R}^2$ and assume that $\mathcal{S}(X) = \mathcal{S}(Y)$. Then for every $\theta \in [0, 1]$ we have*

$$\mathcal{S}(\theta X + (1 - \theta)Y) \le S(X) + \frac{1}{\kappa} \ln(\theta^\kappa + (1 - \theta)^\kappa), \tag{10}$$

where $\kappa > 0$ is a universal constant. We can take $\kappa = 1/5$.

Remark 3. Proving Conjecture 1 is equivalent to showing the above theorem with $\kappa = 1$.

Notice that in the above reverse EPI we estimate the entropy of linear combinations of summands whose joint distribution is log-concave. This is different from what would be the straightforward reverse form of the EPI (3) for independent summands with weights $\sqrt{\lambda}$ and $\sqrt{1 - \lambda}$ preserving variance. Suppose that

the summands X, Y are independent and identically distributed, say with finite variance and recall (4). Then, as we mentioned in the introduction, the EPI says that the function $[0,1] \ni \lambda \to \mathcal{S}(X_\lambda)$ is minimal at $\lambda = 0$ and $\lambda = 1$. Following this logic, reversing the EPI could amount to determining the λ for which the maximum of this function occurs. Our next result shows that the somewhat natural guess of $\lambda = 1/2$ is false in general.

Proposition 2. *For each positive $\lambda_0 < \frac{1}{2(2+\sqrt{2})}$ there is a symmetric continuous random variable X of finite variance for which $\mathcal{S}(X_{\lambda_0}) > \mathcal{S}(X_{1/2})$.*

Nevertheless, we believe that in the log-concave setting the function $\lambda \mapsto \mathcal{S}(X_\lambda)$ should behave nicely.

Conjecture 2. Let X and Y be independent copies of a log-concave random variable. Then the function

$$\lambda \mapsto \mathcal{S}(\sqrt{\lambda}X + \sqrt{1-\lambda}Y)$$

is concave on $[0,1]$.

3 Proofs

3.1 Theorems 1 and 2 are equivalent

To see that Theorem 2 implies Theorem 1 let us take a symmetric log-concave random vector (X, Y) in $\mathbb{R}^2$ and take θ such that $\mathcal{S}(X/\theta) = \mathcal{S}(Y/(1-\theta))$, that is, $\theta = e^{\mathcal{S}(X)}/(e^{\mathcal{S}(X)} + e^{\mathcal{S}(Y)}) \in [0,1]$. Applying Theorem 2 with the vector $(X/\theta, Y/(1-\theta))$ and using the identity $\mathcal{S}(X/\theta) = \mathcal{S}(X) - \ln\theta = -\ln(e^{\mathcal{S}(X)} + e^{\mathcal{S}(Y)})$ gives

$$\mathcal{S}(X + Y) \leq S(X/\theta) + \frac{1}{\kappa}\ln\left(\frac{e^{\kappa\mathcal{S}(X)} + e^{\kappa\mathcal{S}(Y)}}{(e^{\mathcal{S}(X)} + e^{\mathcal{S}(Y)})^\kappa}\right) = \frac{1}{\kappa}\ln\left(e^{\kappa\mathcal{S}(X)} + e^{\kappa\mathcal{S}(Y)}\right),$$

so (9) follows.

To see that Theorem 1 implies Theorem 2, take a log-concave vector (X, Y) with $\mathcal{S}(X) = \mathcal{S}(Y)$ and apply (9) to the vector $(\theta X, (1-\theta)Y)$, which yields

$$\mathcal{S}(\theta X + (1-\theta)Y) \leq \frac{1}{\kappa}\ln\left(\theta^\kappa e^{\kappa\mathcal{S}(X)} + (1-\theta)^\kappa e^{\kappa\mathcal{S}(Y)}\right)$$

$$= \mathcal{S}(X) + \frac{1}{\kappa}\ln\left(\theta^\kappa + (1-\theta)^\kappa\right).$$

3.2 Proof of Remark 2

Let $w\colon \mathbb{R}^2 \longrightarrow [0, +\infty)$ be the density of such a vector and let f, g, h be the densities of $X, Y, X + Y$ as in (7). The assumption means that $f(x) = tg(tx)$. By convexity,

$$\mathcal{S}(X + Y) = \inf\left\{-\int h\ln p, \ p \text{ is a probability density on } \mathbb{R}\right\}.$$

6

Using Fubini's theorem and changing variables yields

$$-\int h \ln p = -\iint w(x, y) \ln p(x + y) \, \mathrm{d}x \mathrm{d}y$$

$$= -\theta(1 - \theta) \iint w(\theta x, (1 - \theta)y) \ln p(\theta x + (1 - \theta)y) \, \mathrm{d}x \mathrm{d}y$$

for every $\theta \in (0, 1)$ and a probability density p. If p is log-concave we get

$$\mathcal{S}(X + Y) \leq -\theta^2(1 - \theta) \iint w(\theta x, (1 - \theta)y) \ln p(x) \, \mathrm{d}x \mathrm{d}y$$

$$-\theta(1 - \theta)^2 \iint w(\theta x, (1 - \theta)y) \ln p(y) \, \mathrm{d}x \mathrm{d}y$$

$$= -\theta^2 \int f(\theta x) \ln p(x) \mathrm{d}x - (1 - \theta)^2 \int g\big((1 - \theta)y\big) \ln p(y) \mathrm{d}y.$$

Set

$$p(x) = \theta f(\theta x) = t\theta g(t\theta x)$$

with θ such that $t\theta = 1 - \theta$. Then the last expression becomes

$$\theta \mathcal{S}(X) + (1 - \theta)\mathcal{S}(Y) - \theta \ln \theta - (1 - \theta) \ln(1 - \theta).$$

Since $\mathcal{S}(Y) = \mathcal{S}(X) + \ln t = \mathcal{S}(X) + \ln \frac{1-\theta}{\theta}$, we thus obtain

$$\mathcal{S}(X + Y) \leq \mathcal{S}(X) - \ln \theta = \mathcal{S}(X) + \ln(1 + t) = \ln \left(e^{\mathcal{S}(X)} + e^{\mathcal{S}(Y)} \right).$$

3.3   Proof of Theorem 2

The idea of our proof of Theorem 2 is very simple. For small θ we bound the quantity $\mathcal{S}(\theta X + (1 - \theta)Y)$ by estimating its derivative. To bound it for large θ, we shall crudely apply Proposition 1. The exact bound based on estimating the derivative reads as follows.

Proposition 3. *Let (X, Y) be a symmetric log-concave random vector on $\mathbb{R}^2$. Assume that $\mathcal{S}(X) = \mathcal{S}(Y)$ and let $0 \leq \theta \leq \frac{1}{2(1+e)}$. Then*

$$S(\theta X + (1 - \theta)Y) \leq S(X) + 60(1 + e)\theta. \tag{11}$$

The main ingredient of the proof of the above proposition is the following lemma. We postpone its proof until the next subsection.

Lemma 1. *Let $w : \mathbb{R}^2 \to \mathbb{R}_+$ be an even log-concave function. Define $f(x) = \int w(x, y) \mathrm{d}y$ and $\gamma = \int w(0, y) \mathrm{d}y \big/ \int w(x, 0) \mathrm{d}x$. Then we have*

$$\iint \frac{-f'(x)}{f(x)} yw(x, y) \mathrm{d}x \mathrm{d}y \leq 30\gamma \int w.$$

Proof of Proposition 3. For $\theta = 0$ both sides of inequality (11) are equal. It is therefore enough to prove that $\frac{d}{d\theta} S(\theta X + (1 - \theta)Y) \leq 60(1 + e)$ for $0 \leq \theta \leq \frac{1}{2(1+e)}$. Let f_θ be the density of $X_\theta = \theta X + (1 - \theta)Y$. Note that $f_\theta = e^{-\varphi_\theta}$, where φ_θ is convex. Let $\frac{d\varphi_\theta}{d\theta} = \Phi_\theta$ and $\frac{df_\theta}{d\theta} = F_\theta$. Then $\Phi_\theta = -F_\theta/f_\theta$. Using the chain rule we get

$$\frac{d}{d\theta} S(\theta X + (1 - \theta)Y) = -\frac{d}{d\theta} \mathbb{E} \ln f_\theta = \frac{d}{d\theta} \mathbb{E} \varphi_\theta(X_\theta)$$
$$= \mathbb{E}\Phi_\theta(X_\theta) + \mathbb{E}\varphi_\theta'(X_\theta)(X - Y).$$

Moreover,

$$\mathbb{E}\Phi_\theta(X_\theta) = -\mathbb{E}F_\theta(X_\theta)/f_\theta(X_\theta) = -\int F_\theta(x)dx$$
$$= -\frac{d}{d\theta}\int f_\theta(x)dx = 0.$$

Let $Z_\theta = (X_\theta, X - Y)$ and let w_θ be the density of Z_θ. Using Lemma 1 with $w = w_\theta$ gives

$$\frac{d}{d\theta} S(\theta X + (1 - \theta)Y) = -\mathbb{E}\left(\frac{f_\theta'(X_\theta)}{f_\theta(X_\theta)}(X - Y)\right)$$
$$= -\int \frac{f_\theta(x)}{f_\theta(x)} y w_\theta(x, y)dxdy \leq 30\gamma_\theta,$$

where $\gamma_\theta = \int w_\theta(0, y)dy / \int w_\theta(x, 0)dx$. It suffices to show that $\gamma_\theta \leq 2(1+e)$ for $0 \leq \theta \leq \frac{1}{2(1+e)}$. Let w be the density of (X, Y). Then $w_\theta(x, y) = w(x + (1 - \theta)y, x - \theta y)$. To finish the proof we again use the fact that $\|v\|_w = (\int w(tv)dt)^{-1}$ is a norm. Note that

$$\gamma_\theta = \frac{\int w_\theta(0, y)dy}{\int w_\theta(x, 0)dx} = \frac{\int w((1 - \theta)y, -\theta y)dy}{\int w(x, x)dx} = \frac{\|e_1 + e_2\|_w}{\|(1 - \theta)e_1 - \theta e_2\|_w}.$$

Let $f(x) = \int w(x, y)dy$ and $g(x) = \int w(y, x)dy$ be the densities of real log-concave random variables X and Y, respectively. Observe that by (6) we have

$$\|f\|_\infty^{-1} \leq e^{S(X)} \leq e\|f\|_\infty^{-1}, \qquad \|g\|_\infty^{-1} \leq e^{S(Y)} \leq e\|g\|_\infty^{-1}.$$

Since $\|f\|_\infty^{-1} = f(0)^{-1} = \|e_1\|_w$, $\|g\|_\infty^{-1} = g(0)^{-1} = \|e_2\|_w$ and $S(X) = S(Y)$, this gives $e^{-1} \leq \|e_1\|/\|e_2\| \leq e$. Thus, by the triangle inequality

$$\gamma_\theta \leq \frac{\|e_1\|_w + \|e_2\|_w}{(1 - \theta)\|e_1\|_w - \theta\|e_2\|_w}$$
$$\leq \frac{(1 + e)\|e_1\|_w}{(1 - \theta)\|e_1\|_w - \theta e\|e_1\|_w} = \frac{1 + e}{1 - \theta(1 + e)}$$
$$\leq 2(1 + e).$$

$\square$

Proof of Theorem 2. We can assume that $\theta \in [0, 1/2]$. Using Proposition 1 with the vector $(\theta X, (1 - \theta)Y)$ and the fact that $\mathcal{S}(X) = \mathcal{S}(Y)$ we get $\mathcal{S}(\theta X + (1 - \theta)Y) \leq \mathcal{S}(X) + 1$. Thus, from Proposition 3 we deduce that it is enough to find $\kappa > 0$ such that

$$\min\{1, 60(1 + e)\theta\} \leq \kappa^{-1} \ln(\theta^\kappa + (1 - \theta)^\kappa), \qquad \theta \in [0, 1/2]$$

(if $60(1 + e)\theta < 1$ then $\theta < \frac{1}{2(1+e)}$ and therefore Proposition 3 indeed can be used in this case). By the concavity and monotonicity of the right hand side it is enough to check this inequality at $\theta_0 = (60(1 + e))^{-1}$, that is, we have to verify the inequality $e^\kappa \leq \theta_0^\kappa + (1 - \theta_0)^\kappa$. We check that this is true for $\kappa = 1/5$. $\square$

3.4 Proof of Lemma 1

We start off by establishing two simple and standard lemmas. The second one is a limiting case of the so-called Grünbaum theorem, see [19] and [26].

Lemma 2. *Let $f : \mathbb{R} \to \mathbb{R}_+$ be an even log-concave function. For $\beta > 0$ define a_β by*

$$a_\beta = \sup\{x > 0, \ f(x) \geq e^{-\beta} f(0)\}.$$

Then we have

$$2e^{-\beta} a_\beta \leq \ \frac{1}{f(0)} \int f \ \leq 2(1 + \beta^{-1} e^{-\beta}) a_\beta.$$

Proof. Since f is even and log-concave, it is maximal at zero and nonincreasing on $[0, \infty)$. Consequently, the left hand inequality immediately follows from the definition of a_β. By comparing $\ln f$ with an appropriate linear function, log-concavity also guarantees that $f(x) \leq f(0)e^{-\beta \frac{x}{a_\beta}}$ for $|x| > a_\beta$, hence

$$\int f \leq 2a_\beta f(0) + 2 \int_{a_\beta}^{\infty} f(0)e^{-\beta \frac{x}{a_\beta}} \, \mathrm{d}x = 2a_\beta f(0) + 2f(0)\frac{a_\beta}{\beta} e^{-\beta}$$

which gives the right hand inequality. $\square$

Lemma 3. *Let X be a log-concave random variable. Let a satisfy $\mathbb{P}(X > a) \leq e^{-1}$. Then $\mathbb{E}X \leq a$.*

Proof. Without loss of generality assume that X is a continuous random variable and that $\mathbb{P}(X > a) = e^{-1}$. Moreover, the statement is translation invariant, so we can assume that $a = 0$. Let $e^{-\varphi}$ be the density of X, where φ is convex. There exists a function ψ of the form

$$\psi(x) = \begin{cases} ax + b, & x \geq L \\ +\infty, & x < L \end{cases}$$

such that $\psi(0) = \varphi(0)$ and $e^{-\psi}$ is the probability density of a random variable Y with $\mathbb{P}(Y > a) = e^{-1}$. One can check, using convexity of φ, that $\mathbb{E}X \leq \mathbb{E}Y$. We have $1 = \int e^{-\psi} = \frac{1}{a}e^{-(b+aL)}$ and $e^{-1} = \int_0^\infty e^{-\psi} = \frac{1}{a}e^{-b}$. It follows that $aL = -1$ and we have $\mathbb{E}X \leq \mathbb{E}Y = \frac{1}{a}\left(L + \frac{1}{a}\right)e^{-(b+aL)} = 0$. $\qquad\square$

We are ready to prove Lemma 1.

Proof of Lemma 1. Without loss of generality let us assume that w is strictly log-concave and $w(0) = 1$. First we derive a pointwise estimate on w which will enable us to obtain good pointwise bounds on the quantity $\int yw(x, y)\mathrm{d}y$, relative to $f(x)$. To this end, set unique positive parameters a and b to be such that $w(a, 0) = e^{-1} = w(0, b)$. Consider $l \in (0, a)$. We have

$$w(-l, 0) = w(l, 0) \geq w(a, 0)^{l/a}w(0, 0)^{1-l/a} = e^{-l/a}.$$

Fix $x > 0$ and let $y > \frac{b}{a}x + b$. Let l be such that the line passing through the points $(0, b)$ and (x, y) intersect the x-axis at $(-l, 0)$, that is $l = \frac{bx}{y-b}$. Note that $l \in (0, a)$. Then

$$e^{-1} = w(0, b) \geq w(x, y)^{b/y}w(-l, 0)^{1-b/y} \geq w(x, y)^{b/y}e^{-\frac{l}{a}(1-b/y)}$$

$$= \left[w(x, y)e^{-\frac{l}{a}\frac{y}{b}\frac{y-b}{y}}\right]^{b/y},$$

hence

$$w(x, y) \leq e^{x/a - y/b}, \quad \text{for } x > 0 \text{ and } y > \frac{b}{a}x + b.$$

Let X be a random variable with log-concave density $y \mapsto w(x, y)/f(x)$. Let us take $\beta = b + b\ln(\max\{f(0), b\})$ and

$$\alpha = \frac{b}{a}x - b\ln f(x) + \beta.$$

Since f is maximal at zero (as it is an even log-concave function), we check that

$$\alpha \geq \frac{b}{a}x - b\ln f(0) + \beta \geq \frac{b}{a}x + b,$$

so we can use the pointwise estimate on w and get

$$\int_\alpha^\infty w(x, y)\mathrm{d}y \leq e^{x/a}\int_\alpha^\infty e^{-y/b}\mathrm{d}y = be^{x/a - \alpha/b} = \frac{b}{\max\{f(0), b\}}e^{-1}f(x) \leq e^{-1}f(x).$$

This means that $\mathbb{P}(X > \alpha) \leq e^{-1}$, which in view of Lemma 3 yields

$$\frac{1}{f(x)}\int yw(x, y)\mathrm{d}y = \mathbb{E}X \leq \alpha = \frac{b}{a}x - b\ln f(x) + \beta, \quad \text{for } x > 0.$$

Having obtained this bound, we can easily estimate the quantity stated in the lemma. By the symmetry of w we have

$$\iint \frac{-f'(x)}{f(x)}yw(x, y)\mathrm{d}x\mathrm{d}y = 2\iint_{x>0}\frac{-f'(x)}{f(x)}yw(x, y)\mathrm{d}x\mathrm{d}y.$$

Since f decreases on $[0, \infty)$, the factor $-f'(x)$ is nonnegative for $x > 0$, thus we can further write

$$\iint \frac{-f'(x)}{f(x)} yw(x,y)\mathrm{d}x\mathrm{d}y \leq 2\int_0^\infty -f'(x)\left(\frac{b}{a}x - b\ln f(x) + \beta\right)\mathrm{d}x$$

$$= 2f(0)(-b\ln f(0) + \beta) + 2\int_0^\infty f(x)\left(\frac{b}{a} - b\frac{f'(x)}{f(x)}\right)\mathrm{d}x$$

$$= 2f(0)b\left(1 + \ln\frac{\max\{f(0),b\}}{f(0)}\right) + \frac{b}{a}\int w + 2f(0)b.$$

Now we only need to put the finishing touches to this expression. By Lemma 2 applied to the functions $x \mapsto w(x,0)$ and $y \mapsto w(0,y)$ we obtain

$$\frac{b}{a} \leq \frac{e}{2}2(1+e^{-1})\frac{\int w(0,y)\mathrm{d}y}{\int w(x,0)\mathrm{d}x} = (e+1)\gamma$$

and $b/f(0) \leq e/2$. Estimating the logarithm yields

$$1 + \ln\frac{\max\{f(0),b\}}{f(0)} \leq \frac{\max\{f(0),b\}}{f(0)} \leq \frac{e}{2}.$$

Finally, by log-concavity,

$$\int w(x,y)\mathrm{d}x\mathrm{d}y \geq \int \sqrt{w(2x,0)w(0,2y)}\mathrm{d}x\mathrm{d}y = \frac{1}{4}\int \sqrt{w(x,0)}\mathrm{d}x \int \sqrt{w(0,y)}\mathrm{d}y$$

and

$$\int w(x,0)\mathrm{d}x \leq \sqrt{w(0,0)}\int \sqrt{w(x,0)}\mathrm{d}x = \int \sqrt{w(x,0)}\mathrm{d}x.$$

Combining these two estimates we get

$$f(0) = \int w(0,y)\mathrm{d}y \leq \int \sqrt{w(0,y)}\mathrm{d}y \leq \frac{4\int w}{\int w(x,0)\mathrm{d}x}$$

and consequently,

$$f(0)b \leq \frac{e}{2}f(0)f(0) \leq 2ef(0)\frac{\int w}{\int w(x,0)\mathrm{d}x} = 2e\gamma\int w.$$

Finally,

$$\iint \frac{-f'(x)}{f(x)} yw(x,y)\mathrm{d}x\mathrm{d}y \leq (2e^2 + 5e + 1)\gamma\int w$$

and the assertion follows. $\qquad\square$

3.5   Proof of Proposition 2

For a real number s and nonnegative numbers $\alpha \leq \beta$ we define the following trapezoidal function

$$T_{\alpha,\beta}^s(x) = \begin{cases} 0 & \text{if } x < s \text{ or } x > s + \alpha + \beta, \\ x - s & \text{if } s \leq x \leq s + \alpha, \\ \alpha & \text{if } s + \alpha \leq x \leq s + \beta, \\ s + \alpha + \beta - x & \text{if } s + \beta \leq x \leq s + \alpha + \beta. \end{cases}$$

The motivation is the following convolution identity: for real numbers a, a' and nonnegative numbers h, h' such that $h \le h'$ we have

$$\mathbf{1}_{[a,a+h]} \star \mathbf{1}_{[a',a'+h']} = T_{h,h'}^{a+a'}. \tag{12}$$

It is also easy to check that

$$\int_{\mathbb{R}} T_{\alpha,\beta}^s = \alpha\beta. \tag{13}$$

We shall need one more formula: for any real number s and nonnegative numbers A, α, β with $\alpha \le \beta$ we have

$$I(A, \alpha, \beta) = \int_{\mathbb{R}} A T_{\alpha,\beta}^s \ln\left(A T_{\alpha,\beta}^s\right) = A\alpha\beta \ln(A\alpha) - \frac{1}{2} A\alpha^2. \tag{14}$$

Fix $0 < a < b = a + h$. Let X be a random variable with the density

$$f(x) = \frac{1}{2h} \left(\mathbf{1}_{[-b,-a]}(x) + \mathbf{1}_{[a,b]}(x)\right).$$

We shall compute the density f_λ of X_λ. Denote $u = \sqrt{\lambda}$, $v = \sqrt{1-\lambda}$ and without loss of generality assume that $\lambda \le 1/2$. Clearly, $f_\lambda(x) = \frac{1}{u} f\left(\frac{\cdot}{u}\right) \star \frac{1}{v} f\left(\frac{\cdot}{v}\right)(x)$, so by (12) we have

$$f_\lambda(x) = \left(\mathbf{1}_{u[-b,-a]} \star \mathbf{1}_{v[-b,-a]} + \mathbf{1}_{u[a,b]} \star \mathbf{1}_{v[-b,-a]}\right.$$
$$\left. + \mathbf{1}_{u[-b,-a]} \star \mathbf{1}_{v[a,b]} + \mathbf{1}_{u[a,b]} \star \mathbf{1}_{v[a,b]}\right)(x) \cdot \frac{1}{(2h)^2 uv}$$
$$= \left(\underbrace{T_{uh,vh}^{-(u+v)b}(x)}_{T_1} + \underbrace{T_{uh,vh}^{ua-vb}(x)}_{T_2} + \underbrace{T_{uh,vh}^{-ub+va}(x)}_{T_3} + \underbrace{T_{uh,vh}^{(u+v)a}(x)}_{T_4}\right) \cdot \frac{1}{(2h)^2 uv}.$$

This symmetric density is superposition of 4 trapezoid functions T_1, T_2, T_3, T_4 which are certain shifts of the same trapezoid function $T_0 = T_{uh,vh}^0$. The shifts may overlap depending on the value of λ. Now we shall consider two particular values of λ.

Case 1: $\lambda = 1/2$. Then $u = v = 1/\sqrt{2}$. Notice that T_0 becomes a triangle looking function and $T_2 = T_3$, so we obtain

$$f_{1/2}(x) = \frac{1}{2h^2} \left(T_{h/\sqrt{2},h/\sqrt{2}}^{-b\sqrt{2}} + 2 T_{h/\sqrt{2},h/\sqrt{2}}^{-h/\sqrt{2}} + T_{h/\sqrt{2},h/\sqrt{2}}^{a\sqrt{2}}\right)(x).$$

If $h/\sqrt{2} < a\sqrt{2}$ then the supports of the summands are disjoint and with the aid of identity (14) we obtain

$$\mathcal{S}(X_{1/2}) = -2I\left(\frac{1}{2h^2}, \frac{h}{\sqrt{2}}, \frac{h}{\sqrt{2}}\right) - I\left(\frac{1}{h^2}, \frac{h}{\sqrt{2}}, \frac{h}{\sqrt{2}}\right) = \ln(2h) + \frac{1}{2}.$$

Case 2: small λ. Now we choose $\lambda = \lambda_0$ so that the supports of T_1 and T_2 intersect in such a way that the down-slope of T_1 adds up to the up-slope of T_2 giving a flat piece. This happens when $-b(u+v) + vh = ua - bv$, that is,

$$\sqrt{\frac{1-\lambda_0}{\lambda_0}} = \frac{v}{u} = \frac{a+b}{h} = 2\frac{a}{h} + 1. \tag{15}$$

The earlier condition $a/h > 1/2$ implies that $\lambda_0 < 1/5$. With the above choice for λ we have $T_1 + T_2 = T_{uh,2vh}^{-b(u+v)}$, hence by symmetry

$$f_\lambda = \left(T_{uh,2vh}^{-b(u+v)} + T_{uh,2vh}^{-ub+va} \right) \cdot \frac{1}{(2h)^2 uv}.$$

As long as $-ub + va > 0$, the supports of these two trapezoid functions are disjoint. Given our choice for λ, this is equivalent to $v/u > b/a = 1 + h/a = 1 + 2/(v/u - 1)$, or putting $v/u = \sqrt{1/\lambda_0 - 1}$, to $\lambda_0 < \frac{1}{2(2+\sqrt{2})}$. Then also $\lambda_0 < 1/5$ and we get

$$\mathcal{S}(X_\lambda) = -2I\left(\frac{1}{(2h)^2 uv}, uh, 2vh \right) = \ln(4vh) + \frac{u}{4v}$$

$$= \ln(4h\sqrt{1-\lambda_0}) + \frac{1}{4}\sqrt{\frac{\lambda_0}{1-\lambda_0}}.$$

We have

$$\mathcal{S}(X_{\lambda_0}) - \mathcal{S}(X_{1/2}) = \ln 2 - \frac{1}{2} + \ln\sqrt{1-\lambda_0} + \frac{1}{4}\sqrt{\frac{\lambda_0}{1-\lambda_0}}.$$

We check that the right hand side is positive for $\lambda_0 < \frac{1}{2(2+\sqrt{2})}$. Therefore, we have shown that for each such λ_0 there is a choice for the parameters a and h (given by (15)), and hence a random variable X, for which $\mathcal{S}(X_{\lambda_0}) > \mathcal{S}(X_{1/2})$.

References

[1] Artstein, S., Ball, K. M., Barthe, F., Naor, A., On the rate of convergence in the entropic central limit theorem, *Probab. Theory Related Fields* 129 (2004), no. 3, 381–390.

[2] Artstein, S., Ball, K. M., Barthe, F., Naor, A., Solution of Shannon's problem on the monotonicity of entropy, *J. Amer. Math. Soc.* 17 (2004), no. 4, 975–982.

[3] Ball, K., Logarithmically concave functions and sections of convex sets in $\mathbb{R}^n$, *Studia Math.* 88 (1988), 69–84.

[4] Ball, K., Barthe, F., Naor, A., Entropy jumps in the presence of a spectral gap, *Duke Math. J.* 119 (2003), no. 1, 41–63.

[5] Ball, K., Nguyen, V. H., Entropy jumps for isotropic log-concave random vectors and spectral gap, *Studia Math.* 213 (2012), no. 1, 81–96.

[6] Blachman, N. M., The convolution inequality for entropy powers, *IEEE Trans. Information Theory* IT-11 (1965) 267–271.

[7] Bobkov, S. G., Chistyakov, G. P., Entropy power inequality for the Rényi entropy, *IEEE Transactions on Information Theory*, vol. 61 (2015), no. 2, pp. 708–714.

[8] Bobkov, S. G., Chistyakov, G. P., Götze, F., Stability problems in Cramér-type characterization in case of i.i.d. summands, *Theory Probab. Appl.* 57(4), 568–588, 2013.

[9] Bobkov, S. G., Chistyakov, G. P., Götze, F., Stability of Cramérs characterization of the normal laws in information distances, *arXiv:1512.03571*, 2016.

[10] Bobkov, S. G., Madiman, M., Reverse Brunn-Minkowski and reverse entropy power inequalities for convex measures, *J. Funct. Anal.* 262 (2012), no. 7, 3309–3339.

[11] Bobkov, S. G., Madiman, M., On the problem of reversibility of the entropy power inequality. *Limit theorems in probability, statistics and number theory*, 61–74, Springer Proc. Math. Stat., 42, *Springer, Heidelberg*, 2013.

[12] Bobkov, S. G., Madiman, M., The entropy per coordinate of a random vector is highly constrained under convexity conditions, *IEEE Trans. Inform. Theory*, 57 (2011), no. 8, 4940–4954.

[13] Carlen, E. A., Superadditivity of Fisher's information and logarithmic Sobolev inequalities, *J. Funct. Anal.* 101 (1991), no. 1, 194–211.

[14] Carlen, E. A., Soffer, A., Entropy production by block variable summation and central limit theorems, *Comm. Math. Phys.* 140 (1991), no. 2, 339–371.

[15] Costa, M. H. M., A new entropy power inequality, *IEEE Trans. Inform. Theory* 31 (1985), no. 6, 751–760.

[16] Cover, T. M., Zhang, Z., On the maximum entropy of the sum of two dependent random variables. *IEEE Trans. Inform. Theory* 40 (1994), no. 4, 1244–1246.

[17] Dembo, A., Simple proof of the concavity of the entropy power with respect to added Gaussian noise, *IEEE Trans. Inform. Theory* 35 (1989), no. 4, 887–888.

[18] Dembo, A., Cover, T. M., Thomas, J. A., Information-theoretic inequalities, *IEEE Trans. Inform. Theory* 37 (1991), no. 6, 1501–1518.

[19] Grünbaum, B., Partitions of mass-distributions and of convex bodies by hyperplanes, *Pacific J. Math.* 10 (1960), 1257–1261.

[20] Johnson, O., Barron, A., Fisher information inequalities and the central limit theorem, *Probab. Theory Related Fields* 129 (2004), no. 3, 391–409.

[21] Kalton, N. J., Peck, N. T., Roberts, James W., An F-space sampler. London Mathematical Society Lecture Note Series, 89. *Cambridge University Press, Cambridge*, 1984.

[22] Lieb, E. H., Proof of an entropy conjecture of Wehrl, *Comm. Math. Phys.* 62 (1978), no. 1, 35–41.

[23] Mokshay Madiman, M., Kontoyiannis, I., The Ruzsa divergence for random elements in locally compact abelian groups, arXiv:1508.04089.

[24] Shannon, C. E., A mathematical theory of communication, *Bell System Tech. J.* 27, (1948). 379–423, 623–656.

[25] Stam, A. J., Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Information and Control* 2 1959, 101–112.

[26] Szarek, S., On measures of symmetry and floating bodies, arXiv:1302.2076

[27] Verdú, S., Guo, D., A simple proof of the entropy-power inequality, *IEEE Trans. Inform. Theory* 52 (2006), no. 5, 2165–2166.

[28] Villani, C., A short proof of the "concavity of entropy power", *IEEE Trans. Inform. Theory* 46 (2000), no. 4, 1695–1696.

Keith Ball[*], `k.m.ball@warwick.ac.uk`

Piotr Nayar[†], `nayar@mimuw.edu.pl`

Tomasz Tkocz[*], `t.tkocz@warwick.ac.uk`

[*]Mathematics Institute, University of Warwick
Coventry CV4 7AL
UK

[†]Institute of Mathematics & Applications
Minneapolis MN 55455
United States