On the Problem of Reversibility of the Entropy Power Inequality

Sergey G. Bobkov and Mokshay M. Madiman

Dedicated to Friedrich Götze on the occasion of his sixtieth birthday

Abstract As was shown recently by the authors, the entropy power inequality can be reversed for independent summands with sufficiently concave densities, when the distributions of the summands are put in a special position. In this note it is proved that reversibility is impossible over the whole class of convex probability distributions. Related phenomena for identically distributed summands are also discussed.

Keywords Convex measures • Entropy power inequality • Log-concave • Reverse Brunn-Minkowski inequality • Rogers-Shephard inequality

2010 *Mathematics Subject Classification.* 60F05.

S.G. Bobkov (✉)
School of Mathematics, University of Minnesota, Vincent Hall 228, 206 Church St SE, Minneapolis, MN 55455, USA
e-mail: bobkov@math.umn.edu

M.M. Madiman
Department of Mathematical Sciences, University of Delaware, 501 Ewing Hall, Newark, DE 19716, USA
e-mail: mokshay.madiman@yale.edu

P. Eichelsbacher et al. (eds.), *Limit Theorems in Probability, Statistics and Number Theory*, Springer Proceedings in Mathematics & Statistics 42,
DOI 10.1007/978-3-642-36068-8_4,

1 The Reversibility Problem for the Entropy Power Inequality

Given a random vector X in $\mathbb{R}^n$ with density f, introduce the entropy functional (or Shannon's entropy)

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

and the entropy power

$$H(X) = e^{2h(X)/n},$$

provided that the integral exists in the Lebesgue sense. For example, if X is uniformly distributed in a convex body $A \subset \mathbb{R}^n$, we have

$$h(X) = \log |A|, \qquad H(X) = |A|^{2/n},$$

where $|A|$ stands for the n-dimensional volume of A.

The entropy power inequality due to Shannon and Stam indicates that

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

for any two independent random vectors X and Y in $\mathbb{R}^n$, for which the entropy is defined ([27, 28], cf. also [14, 15, 29]). This is one of the fundamental results in Information Theory, and it is of large interest to see how sharp (1) is.

The equality here is only achieved, when X and Y have normal distributions with proportional covariance matrices. Note that the right-hand side is unchanged when X and Y are replaced with affine volume-preserving transformation, that is, with random vectors

$$\tilde{X} = T_1(X), \qquad \tilde{Y} = T_2(Y) \qquad (|\det T_1| = |\det T_2| = 1). \tag{2}$$

On the other hand, the entropy power $H(\tilde{X} + \tilde{Y})$ essentially depends on the choice of T_1 and T_2. Hence, it is reasonable to consider a formally improved variant of (1),

$$\inf_{T_1, T_2} H(\tilde{X} + \tilde{Y}) \geq H(X) + H(Y), \tag{3}$$

where the infimum is running over all affine maps $T_1, T_2 : \mathbb{R}^n \to \mathbb{R}^n$ subject to (2). (Note that one of these maps may be taken to be the identity operator.) Now, equality in (3) is achieved, whenever X and Y have normal distributions with arbitrary positive definite covariance matrices.

A natural question arises: When are both the sides of (3) of a similar order? For example, within a given class of probability distributions (of X and Y), one wonders whether or not it is possible to reverse (3) to get

$$\inf_{T_1, T_2} H(\tilde{X} + \tilde{Y}) \leq C(H(X) + H(Y)) \tag{4}$$

with some constant C.

The question is highly non-trivial already for the class of uniform distributions on convex bodies, when it becomes to be equivalent (with a different constant) to the inverse Brunn-Minkowski inequality

$$\inf_{T_1,T_2} \left|\tilde{A}+\tilde{B}\right|^{1/n} \le C\left(|A|^{1/n}+|B|^{1/n}\right). \tag{5}$$

Here $\tilde{A}+\tilde{B} = \{x+y : x \in \tilde{A},\ y \in \tilde{B}\}$ stands for the Minkowski sum of the images $\tilde{A} = T_1(A)$, $\tilde{B} = T_2(B)$ of arbitrary convex bodies A and B in $\mathbb{R}^n$. To recover such an equivalence, one takes for X and Y independent random vectors uniformly distributed in A and B. Although the distribution of $X+Y$ is not uniform in $A+B$, there is a general entropy-volume relation

$$\frac{1}{4}\,|A+B|^{2/n} \le\ H(X+Y) \le\ |A+B|^{2/n},$$

which may also be applied to the images $\tilde{A}$, $\tilde{B}$ and $\tilde{X}$, $\tilde{Y}$ (cf. [3]).

The inverse Brunn-Minkowski inequality (5) is indeed true and represents a deep result in Convex Geometry discovered by V. D. Milman in the mid 1980s (cf. [21–24]). It has connections with high dimensional phenomena, and we refer an interested reader to [1, 12, 16, 17]. The questions concerning possible description of the maps T_1 and T_2 and related isotropic properties of the normalized Gaussian measures are discussed in [6].

Based on (5), and involving Berwald's inequality in the form of C. Borell [9], the inverse entropy power inequality (4) has been established recently [2,3] for the class of all probability distributions having log-concave densities. Involving additionally a general submodularity property of entropy [19], it turned out also possible to consider more general densities of the form

$$f(x) = V(x)^{-\beta}, \qquad x \in \mathbb{R}^n, \tag{6}$$

where V are positive convex functions on $\mathbb{R}^n$ and $\beta \ge n$ is a given parameter. More precisely, the following statement can be found in [3].

Theorem 1.1. *Let X and Y be independent random vectors in $\mathbb{R}^n$ with densities of the form (6) with $\beta \ge 2n+1$, $\beta \ge \beta_0 n$ ($\beta_0 > 2$). There exist linear volume preserving maps $T_i : \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$H\left(\tilde{X}+\tilde{Y}\right) \le\ C_{\beta_0}\,(H(X)+H(Y)), \tag{7}$$

where $\tilde{X} = T_1(X)$, $\tilde{Y} = T_2(Y)$, and where C_{β_0} is a constant, depending on β_0, only.

The question of what maps T_1 and T_2 can be used in Theorem 1.1 is rather interesting, but certainly the maps that put the distributions of X and Y in M-position suffice (see [3] for terminology and discussion). In a more relaxed form, one needs to have in some sense "similar" positions for both distributions. For example, when considering identically distributed random vectors, there is no need to appeal in Theorem 1.1 to some (not very well understood) affine volume-preserving transformations, since the distributions of X and Y have the same M-ellipsoid. In other words, we have for X and Y drawn independently from the *same* distribution (under the same assumption on form of density as Theorem 1.1) that

$$H(X+Y) \le C_{\beta_0}\,(H(X)+H(Y)) = 2C_{\beta_0}\,H(X). \tag{8}$$

Since the distributions of X and $-Y$ also have the same M-ellipsoid, it is also true that

$$H(X-Y) \le C_{\beta_0}\,(H(X)+H(Y)) = 2C_{\beta_0}\,H(X). \tag{9}$$

We strengthen this observation by providing a quantitative version with explicit constants below (under, however, a convexity condition on the convolved measure). Moreover, one can give a short and relatively elementary proof of it without appealing to Theorem 1.1.

Theorem 1.2. *Let X and Y be independent identically distributed random vectors in $\mathbb{R}^n$ with finite entropy. Suppose that $X-Y$ has a probability density function of the form* (6) *with $\beta \ge \max\{n+1, \beta_0 n\}$ for some fixed $\beta_0 > 1$. Then*

$$H(X-Y) \le D_{\beta_0} H(X)$$

and

$$H(X+Y) \le D_{\beta_0}^2 H(X),$$

where $D_{\beta_0} = \exp(\frac{2\beta_0}{\beta_0-1})$.

In the special case of X and Y being log-concave, a similar quantitative result was recently obtained by [18] using a different approach.

Let us return to Theorem 1.1 and the class of distributions involved there. For growing β, the families (6) shrink and converge in the limit as $\beta \to +\infty$ to the family of log-concave densities which correspond to the class of log-concave probability measures. Through inequalities of the Brunn-Minkowski-type, the latter class was introduced by A. Prékopa [25], while the general case $\beta \ge n$ was studied by C. Borell [10, 11], cf. also [5, 13]. In [10, 11] it was shown that probability measures μ on $\mathbb{R}^n$ with densities (6) (and only they, once μ is absolutely continuous) satisfy the geometric inequality

$$\mu\big(tA+(1-t)B\big) \ge \big[\,t\mu(A)^{\kappa} + (1-t)\mu(B)^{\kappa}\big]^{1/\kappa} \tag{10}$$

for all $t \in (0,1)$ and for all Borel measurable sets $A, B \subset \mathbb{R}^n$, with negative power

$$\kappa = -\frac{1}{\beta-n}.$$

Such μ's form the class of so-called κ-concave measures. In this hierarchy the limit case $\beta = n$ corresponds to $\kappa = -\infty$ and describes the largest class of measures on $\mathbb{R}^n$, called *convex*, in which case (10) turns into

$$\mu(tA + (1-t)B) \geq \min\{\mu(A), \mu(B)\}.$$

This inequality is often viewed as the weakest convexity hypothesis about a given measure μ.

One may naturally wonder whether or not it is possible to relax the assumption on the range of β in (7)–(9), or even to remove any convexity hypotheses. In this note we show that this is impossible already for the class of all one-dimensional convex probability distributions. Note that in dimension one there are only two admissible linear transformations, $\tilde{X} = X$ and $\tilde{X} = -X$, so that one just wants to estimate $H(X+Y)$ or $H(X-Y)$ from above in terms of $H(X)$. As a result, the following statement demonstrates that Theorem 1.1 and its particular cases (8)–(9) are false over the full class of convex measures.

Theorem 1.3. *For any constant C, there is a convex probability distribution μ on the real line with a finite entropy, such that*

$$\min\{H(X+Y), H(X-Y)\} \geq C\, H(X),$$

where X and Y are independent random variables, distributed according to μ.

A main reason for $H(X+Y)$ and $H(X-Y)$ to be much larger than $H(X)$ is that the distributions of the sum $X+Y$ and the difference $X-Y$ may lose convexity properties, when the distribution μ of X is not "sufficiently convex". For example, in terms of the convexity parameter κ (instead of β), the hypothesis of Theorem 1.1 is equivalent to

$$\kappa \geq -\frac{1}{(\beta_0 - 1)n} \quad (\beta_0 > 2), \qquad \kappa \geq -\frac{1}{n+1}.$$

That is, for growing dimension n we require that κ be sufficiently close to zero (or the distributions of X and Y should be close to the class of log-concave measures). These conditions ensure that the convolution of μ with the uniform distribution on a proper (specific) ellipsoid remains to be convex, and its convexity parameter can be controled in terms of β_0 (a fact used in the proof of Theorem 1.1). However, even if κ is close to zero, one cannot guarantee that $X+Y$ or $X-Y$ would have convex distributions.

We prove Theorem 1.2 in Sect. 2 and Theorem 1.3 in Sect. 3, and then conclude in Sect. 4 with remarks on the relationship between Theorem 1.3 and recent results about Cramer's characterization of the normal law.

2 A "Difference Measure" Inequality for Convex Measures

Given two convex bodies A and B in $\mathbb{R}^n$, introduce $A - B = \{x - y : x \in A,\ y \in B\}$. In particular, $A - A$ is called the "difference body" of A. Note it is always symmetric about the origin.

The Rogers-Shephard inequality [26] states that, for any convex body $A \subset \mathbb{R}^n$,

$$|A - A| \ \le\ C_{2n}^n\, |A|, \tag{11}$$

where $C_n^k = \frac{n!}{k!(n-k)!}$ denote usual combinatorial coefficients. Observe that putting the Brunn-Minkowski inequality and (11) together immediately yields that

$$2 \le \frac{|A-A|^{\frac{1}{n}}}{|A|^{\frac{1}{n}}} \le \left[C_{2n}^n\right]^{\frac{1}{n}} < 4,$$

which constrains severely the volume radius of the difference body of A relative to that of A itself. In analogy to the Rogers-Shephard inequality, we ask the following question for entropy of convex measures.

Question. *Let X and Y be independent random vectors in $\mathbb{R}^n$, which are identically distributed with density $V^{-\beta}$, with V positive convex, and $\beta \ge n+\gamma$. For what range of $\gamma > 0$ is it true that $H(X - Y) \le C_\gamma H(X)$, for some constant C_γ depending only on γ?*

Theorems 1.2 and 1.3 partially answer this question. To prove the former, we need the following lemma about convex measures, proved in [4].

Lemma 2.1. *Fix $\beta_0 > 1$. Assume a random vector X in $\mathbb{R}^n$ has a density $f = V^{-\beta}$, where V is a positive convex function on the supporting set. If $\beta \ge n + 1$ and $\beta \ge \beta_0 n$, then*

$$\log\, \|f\|_\infty^{-1} \le\ h(X) \le\ c_{\beta_0} n + \log\, \|f\|_\infty^{-1}, \tag{12}$$

where one can take for the constant $c_{\beta_0} = \frac{\beta_0}{\beta_0 - 1}$.

In other words, for sufficiently convex probability measures, the entropy may be related to the L^∞-norm $\|f\|_\infty = \sup_x f(x)$ of the density f (which is necessarily finite). Observe that the left inequality in (12) is general: It trivially holds without any convexity assumption. On the other hand, the right inequality is an asymptotic version of a result from [4] about extremal role of the multidimensional Pareto distributions.

Now, let f denote the density of the random variable $W = X - Y$ in Theorem 1.2. It is symmetric (even) and thus maximized at zero, by the convexity hypothesis. Hence, by Lemma 2.1,

$$h(W) \le \log \|f\|_\infty^{-1} + c_{\beta_0} n = \log f(0)^{-1} + c_{\beta_0} n.$$

But, if p is the density of X, then $f(0) = \int_{\mathbb{R}^n} p(x)^2\,dx$, and hence

$$\log f(0)^{-1} = -\log \int_{\mathbb{R}^n} p(x) \cdot p(x)\,dx \leq \int_{\mathbb{R}^n} p(x)[-\log p(x)]\,dx$$

by using Jensen's inequality. Combining the above two displays immediately yields the first part of Theorem 1.2.

To obtain the second part, we need the following lemma on the submodularity of the entropy of sums proved in [19].

Lemma 2.2. *Given independent random vectors X, Y, Z in $\mathbb{R}^n$ with absolutely continuous distributions, we have*

$$h(X+Y+Z) + h(Z) \leq h(X+Z) + h(Y+Z),$$

provided that all entropies are well-defined and finite.

Taking X, Y and $-Z$ to be identically distributed, and using the monotonicity of entropy (after adding an independent summand), we obtain

$$h(X+Y) + h(Z) \leq h(X+Y+Z) + h(Z) \leq h(X+Z) + h(Y+Z)$$

and hence

$$h(X+Y) + h(X) \leq 2h(X-Y).$$

Combining this bound with the first part of Theorem 1.2 immediately gives the second part.

It would be more natural to state Theorem 1.2 under a shape condition on the distribution of X rather than on that of $X - Y$, but for this we need to have better understanding of the convexity parameter of the convolution of two κ-concave measures when $\kappa < 0$.

Observe that in the log-concave case of Theorem 1.2 (which is the case of $\beta \to \infty$, but can easily be directly derived in the same way without taking a limit), one can impose only a condition on the distribution of X (rather than that of $X-Y$) since closedness under convolution is guaranteed by the Prékopa-Leindler inequality.

Corollary 2.3. *Let X and Y be independent random vectors in $\mathbb{R}^n$ with log-concave densities. Then*

$$h(X-Y) \leq h(X) + n,$$
$$h(X+Y) \leq h(X) + 2n.$$

In particular, observe that putting the entropy power inequality (1) and Corollary 2.3 together immediately yields that

$$2 \le \frac{H(X-Y)}{H(X)} \le e^2,$$

which constrains severely the entropy power of the "difference measure" of μ relative to that of μ itself.

A result similar to Corollary 2.3 (but with different constants) was recently obtained in [18] using a different approach.

3 Proof of Theorem 1.3

Given a (large) parameter $b > 1$, let a random variable X_b have a truncated Pareto distribution μ, namely, with the density

$$f(x) = \frac{1}{x \log b}\, 1_{\{1<x<b\}}(x).$$

By the construction, μ is supported on a bounded interval $(1, b)$ and is convex.

First we are going to test the inequality

$$H(X_b + Y_b) \le C H(X_b) \tag{13}$$

for growing b, where Y_b is an independent copy of X_b. Note that

$$\begin{aligned} h(X_b) &= \int_1^b f(x)\, \log(x \log b)\, dx \\ &= \log\log b + \frac{1}{\log b} \int_1^b \frac{\log x}{x}\, dx \ = \ \log\log b + \frac{1}{2}\, \log b, \end{aligned}$$

so $H(X_b) = b\, \log^2 b$.

Now, let us compute the convolution of f with itself. The sum $X_b + Y_b$ takes values in the interval $(2, 2b)$. Given $2 < x < 2b$, we have

$$g(x) = (f * f)(x) = \int_{-\infty}^{+\infty} f(x-y) f(y)\, dy = \frac{1}{\log^2 b} \int_\alpha^\beta \frac{dy}{(x-y)y},$$

where the limits of integration are determined to satisfy the constraints $1 < y < b$, $1 < x - y < b$. So,

$$\alpha = \max(1, x-b), \quad \beta = \min(b, x-1),$$

and using $\frac{1}{(x-y)y} = \frac{1}{x}\,(\frac{1}{y} + \frac{1}{x-y})$, we find that

$$g(x) = \frac{1}{x\log^2 b}\,\big(\log(y) - \log(x-y)\big)\Big|_{x=\alpha}^{\beta} = \frac{1}{x\log^2 b}\,\log\frac{y}{x-y}\,\Big|_{x=\alpha}^{\beta}$$
$$= \frac{1}{x\log^2 b}\left(\log\frac{\beta}{x-\beta} - \log\frac{\alpha}{x-\alpha}\right).$$

Note that $x - \alpha = x - \max(1, x-b) = \min(b, x-1) = \beta$. Hence,

$$g(x) = \frac{2}{x\log^2 b}\,\log\frac{\beta}{\alpha} = \frac{2}{x\log^2 b}\,\log\frac{\min(b, x-1)}{\max(1, x-b)}.$$

Equivalently,

$$g(x) = \frac{2}{x\log^2 b}\,\log(x-1), \text{ for } 2 < x < b+1,$$
$$g(x) = \frac{2}{x\log^2 b}\,\log\frac{b}{x-b}, \text{ for } b+1 < x < 2b.$$

Now, on the second interval $b+1 < x < 2b$, we have

$$g(x) \le \frac{2}{x\log^2 b}\,\log b = \frac{2}{x\log b} < \frac{2}{(b+1)\log b} < 1,$$

where the last bound holds for $b \ge e$, for example. Similarly, on the first interval $2 < x < b+1$, using $\log(x-1) < \log b$, we get

$$g(x) \le \frac{2}{x\log b} < \frac{1}{\log b} \le 1.$$

Thus, as soon as $b \ge e$, we have $g \le 1$ on the support interval. From this,

$$h(X_b + Y_b) = \int_2^{2b} g(x)\log(1/g(x))\,dx \ge \int_2^{b} g(x)\log(1/g(x))\,dx.$$

Next, using on the first interval the bound $g(x) \le \frac{2}{x\log b} \le \frac{1}{x}$, valid for $b \ge e^2$, we get for such values of b that

$$h(X_b + Y_b) \ge \int_2^{b} g(x)\log x\,dx = \frac{2}{\log^2 b}\int_2^{b} \frac{\log(x-1)\,\log x}{x}\,dx.$$

To further simplify, we may write $x - 1 \ge \frac{x}{2}$, which gives

$$\int_2^b \frac{\log(x-1)\,\log x}{x}\,dx \geq \int_2^b \frac{\log^2 x}{x}\,dx - \log 2 \int_2^b \frac{\log x}{x}\,dx$$
$$= \frac{1}{3}\left(\log^3 b - \log^3 2\right) - \frac{\log 2}{2}\left(\log^2 b - \log^2 2\right)$$
$$> \frac{1}{3}\,\log^3 b - \frac{\log 2}{2}\,\log^2 b.$$

Hence, $h(X_b + Y_b) > \frac{2}{3}\,\log b - \log 2$, and so

$$H(X_b + Y_b) > \frac{1}{4} b^{4/3} \quad (b \geq e^2).$$

In particular,

$$\frac{H(X_b + Y_b)}{H(X_b)} > \frac{b^{1/3}}{4 \log^2 b} \to +\infty, \quad \text{as } b \to +\infty.$$

Hence, the inequality (13) may not hold for large b with any prescribed value of C.

To test the second bound

$$H(X_b - Y_b) \leq C H(X_b), \tag{14}$$

one may use the previous construction. The random variable $X_b - Y_b$ can take any value in the interval $|x| < b - 1$, where it is described by the density

$$h(x) = \int_{-\infty}^{+\infty} f(x+y) f(y)\,dy = \frac{1}{\log^2 b} \int_\alpha^\beta \frac{dy}{(x+y)y}.$$

Here the limits of integration are determined to satisfy $1 < y < b$ and $1 < x + y < b$. So, assuming for simplicity that $0 < x < b - 1$, the limits are

$$\alpha = 1, \quad \beta = b - x.$$

Writing $\frac{1}{(x+y)y} = \frac{1}{x}\,(\frac{1}{y} - \frac{1}{x+y})$, we find that

$$h(x) = \frac{1}{x \log^2 b}\left(\log(y) - \log(x+y)\right|_{x=\alpha}^{\beta} = \frac{1}{x \log^2 b}\,\log\frac{(b-x)(x+1)}{b}.$$

It should also be clear that

$$h(0) = \frac{1}{\log^2 b} \int_1^b \frac{dy}{y^2} = \frac{1 - \frac{1}{b}}{\log^2 b}.$$

Using $\log \frac{(b-x)(x+1)}{b} < \log(x+1) < x$, we obtain that $h(x) < \frac{1}{\log^2 b} \leq 1$, for $b \geq e^2$.

In this range, since $\frac{(b-x)(x+1)}{b} < b$, we also have that $h(x) \leq \frac{1}{x \log b} \leq \frac{1}{x}$. Hence, in view of the symmetry of the distribution of $X_b - Y_b$,

$$\begin{aligned} h(X_b - Y_b) &= 2 \int_0^{b-1} h(x) \log(1/h(x))\, dx \\ &\geq 2 \int_0^{b/2} h(x) \log x\, dx \\ &= \frac{2}{\log^2 b} \int_2^{b/2} \frac{\log x}{x} \log \frac{(b-x)(x+1)}{b}\, dx. \end{aligned}$$

But for $0 < x < b/2$,

$$\log \frac{(b-x)(x+1)}{b} > \log \frac{x+1}{2} > \log x - \log 2,$$

so

$$\begin{aligned} h(X_b - Y_b) &> \frac{2}{\log^2 b} \int_2^{b/2} \frac{\log^2 x - \log 2 \log x}{x}\, dx \\ &= \frac{2}{\log^2 b} \left(\frac{1}{3} (\log^3(b/2) - \log^3 2) - \frac{\log 2}{2} (\log^2(b/2) - \log^2 2) \right) \\ &> \frac{2}{\log^2 b} \left(\frac{1}{3} \log^3(b/2) - \frac{1}{2} \log^2(b/2) \right) \\ &\sim \frac{2}{3} \log b. \end{aligned}$$

Therefore, like on the previous step, $H(X_b - Y_b)$ is bounded from below by a function, which is equivalent to $b^{4/3}$. Thus, for large b, the inequality (14) may not hold either.

Theorem 1.3 is proved.

4 Remarks

For a random variable X having a density, consider the entropic distance from the distribution of X to normality

$$D(X) = h(Z) - h(X),$$

where Z is a normal random variable with parameters $\mathbb{E}Z = \mathbb{E}X$, $\mathrm{Var}(Z) = \mathrm{Var}(X)$. This functional is well-defined for the class of all probability distributions on the line with finite second moment, and in general $0 \leq D(X) \leq +\infty$.

The entropy power inequality implies that

$$D(X+Y) \leq \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\, D(X) + \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\, D(X) \leq \max(D(X), D(Y)), \tag{15}$$

where $\sigma_1^2 = \mathrm{Var}(X)$, $\sigma_2^2 = \mathrm{Var}(Y)$.

In turn, if X and Y are identically distributed, then Theorem 1.3 reads as follows: For any positive constant c, there exists a convex probability measure μ on $\mathbb{R}$ with X, Y independently distributed according to μ, with

$$D(X \pm Y) \leq D(X) - c.$$

This may be viewed as a strengthened variant of (15). That is, in Theorem 1.3 we needed to show that both $D(X+Y)$ and $D(X-Y)$ may be much smaller than $D(X)$ in the additive sense. In particular, $D(X)$ has to be very large when c is large. For example, in our construction of the previous section

$$\mathbb{E}X_b = \frac{b-1}{\log b}, \quad \mathbb{E}X_b^2 = \frac{b^2-1}{2\log b},$$

which yields

$$D(X_b) \sim \frac{3}{2}\log b, \quad D(X_b+Y_b) \sim \frac{4}{3}\log b,$$

as $b \to +\infty$.

In [7, 8] a slightly different question, raised by M. Kac and H. P. McKean [20] (with the desire to quantify in terms of entropy the Cramer characterization of the normal law), has been answered. Namely, it was shown that $D(X+Y)$ may be as small as we wish, while $D(X)$ is separated from zero. In the examples of [8], $D(X)$ is of order 1, while for Theorem 1.3 it was necessary to use large values for $D(X)$, arbitrarily close to infinity. In addition, the distributions in [7, 8] are not convex.

Acknowledgements Sergey G. Bobkov was supported in part by the NSF grant DMS-1106530. Mokshay M. Madiman was supported in part by the NSF CAREER grant DMS-1056996.

References

1. S. Artstein-Avidan, V. Milman, Y. Ostrover, The M-ellipsoid, symplectic capacities and volume. Comment. Math. Helv. **83**(2), 359–369 (2008)
2. S. Bobkov, M. Madiman, Dimensional behaviour of entropy and information. C. R. Acad. Sci. Paris Sér. I Math. **349**, 201–204 (2011)
3. S. Bobkov, M. Madiman, Reverse Brunn-Minkowski and reverse entropy power inequalities for convex measures. J. Funct. Anal. **262**(7), 3309–3339 (2012)
4. S. Bobkov, M. Madiman, The entropy per coordinate of a random vector is highly constrained under convexity conditions. IEEE Trans. Inform. Theor. **57**(8), 4940–4954 (2011)
5. S.G. Bobkov, Large deviations and isoperimetry over convex probability measures. Electron. J. Probab. **12**, 1072–1100 (2007)
6. S.G. Bobkov, On Milman's ellipsoids and M-position of convex bodies, in *Concentration, Functional Inequalities and Isoperimetry*. Contemporary Mathematics, vol. 545 (American Mathematical Society, Providence, 2011), pp. 23–34
7. S.G. Bobkov, G.P. Chistyakov, F. Götze, Entropic instability of Cramer's characterization of the normal law, in *Selected Works of Willem van Zwet*. Sel. Works Probab. Stat. (Springer, New York, 2012), pp. 231–242 (2012)
8. S.G. Bobkov, G.P. Chistyakov, F. Götze, Stability problems in Cramér-type characterization in case of i.i.d. summands. Teor. Veroyatnost. i Primenen. **57**(4), 701–723 (2011)
9. C. Borell, Complements of Lyapunov's inequality. Math. Ann. **205**, 323–331 (1973)
10. C. Borell, Convex measures on locally convex spaces. Ark. Math. **12**, 239–252 (1974)
11. C. Borell, Convex set functions in d-space. Period. Math. Hungar. **6**(2), 111–136 (1975)
12. J. Bourgain, B. Klartag, V.D. Milman, Symmetrization and isotropic constants of convex bodies, in *Geometric Aspects of Functional Analysis (1986/87)*. Lecture Notes in Mathematics, vol. 1850 (Springer, Berlin, 2004), pp. 101–115
13. H.J. Brascamp, E.H. Lieb, On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. J. Funct. Anal. **22**(4), 366–389 (1976)
14. M. Costa, T.M. Cover, On the similarity of the entropy power inequality and the Brunn-Minkowski inequality. IEEE Trans. Inform. Theor. **IT-30**, 837–839 (1984)
15. A. Dembo, T. Cover, J. Thomas, Information-theoretic inequalities. IEEE Trans. Inform. Theor. **37**(6), 1501–1518 (1991)
16. B. Klartag, V.D. Milman, Geometry of log-concave functions and measures. Geom. Dedicata **112**, 169–182 (2005)
17. H. Koenig, N. Tomczak-Jaegermann, Geometric inequalities for a class of exponential measures. Proc. Am. Math. Soc. **133**(4), 1213–1221 (2005)
18. M. Madiman, I. Kontoyiannis, The Ruzsa divergence for random elements in locally compact abelian groups. Preprint (2013)
19. M. Madiman, On the entropy of sums, in *Proceedings of the IEEE Information Theory Workshop*, Porto, Portugal (IEEE, New York, 2008)
20. H.P. McKean Jr., Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas. Arch. Rational Mech. Anal. **21**, 343–367 (1966)
21. V.D. Milman, An inverse form of the Brunn-Minkowski inequality, with applications to the local theory of normed spaces. C. R. Acad. Sci. Paris Sér. I Math. **302**(1), 25–28 (1986)
22. V.D. Milman, Entropy point of view on some geometric inequalities. C. R. Acad. Sci. Paris Sér. I Math. **306**(14), 611–615 (1988)
23. V.D. Milman, Isomorphic symmetrizations and geometric inequalities, in *Geometric Aspects of Functional Analysis (1986/87)*. Lecture Notes in Mathematics, vol. 1317 (Springer, Berlin, 1988), pp. 107–131
24. G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*. Cambridge Tracts in Mathematics, vol. 94 (Cambridge University Press, Cambridge, 1989)

25. A. Prékopa, Logarithmic concave measures with applications to stochastic programming. Acta Sci. Math. Szeged **32**, 301–316 (1971)
26. C.A. Rogers, G.C. Shephard, The difference body of a convex body. Arch. Math. (Basel) **8**, 220–233 (1957)
27. C.E. Shannon, A mathematical theory of communication. Bell System Tech. J. **27**, 379–423, 623–656 (1948)
28. A.J. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon. Inform. Contr. **2**, 101–112 (1959)
29. S. Szarek, D. Voiculescu, Shannon's entropy power inequality via restricted Minkowski sums, in *Geometric Aspects of Functional Analysis*. Lecture Notes in Mathematics, vol. 1745 (Springer, Berlin, 2000), pp. 257–262