Chapter 4
Further Investigations of Rényi Entropy Power Inequalities and an Entropic Characterization of s-Concave Densities

Jiange Li, Arnaud Marsiglietti, and James Melbourne

Abstract We investigate the role of convexity in Rényi entropy power inequalities. After proving that a general Rényi entropy power inequality in the style of Bobkov and Chistyakov (IEEE Trans Inform Theory 61(2):708–714, 2015) fails when the Rényi parameter $r \in (0, 1)$, we show that random vectors with s-concave densities do satisfy such a Rényi entropy power inequality. Along the way, we establish the convergence in the Central Limit Theorem for Rényi entropies of order $r \in (0, 1)$ for log-concave densities and for compactly supported, spherically symmetric and unimodal densities, complementing a celebrated result of Barron (Ann Probab 14:336–342, 1986). Additionally, we give an entropic characterization of the class of s-concave densities, which extends a classical result of Cover and Zhang (IEEE Trans Inform Theory 40(4):1244–1246, 1994).

4.1 Introduction

Let X be a random vector in $\mathbb{R}^d$. Suppose that X has the density f with respect to the Lebesgue measure. For $r \in (0, 1) \cup (1, \infty)$, the Rényi entropy of order r (or simply, r-Rényi entropy) is defined as

$$h_r(X) = \frac{1}{1-r} \log \int_{\mathbb{R}^d} f(x)^r dx. \tag{4.1}$$

J. Li
Einstein Institute of Mathematics, Hebrew University of Jerusalem, Jerusalem, Israel
e-mail: jiange.li@mail.huji.ac.il

A. Marsiglietti (✉)
Department of Mathematics, University of Florida, Gainesville, FL, USA
e-mail: a.marsiglietti@ufl.edu

J. Melbourne
Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA
e-mail: melbo013@umn.edu

B. Klartag, E. Milman (eds.), *Geometric Aspects of Functional Analysis*,
Lecture Notes in Mathematics 2266,
https://doi.org/10.1007/978-3-030-46762-3_4

For $r \in \{0, 1, \infty\}$, the r-Rényi entropy can be extended continuously such that the RHS of (4.1) is $\log|\mathrm{supp}(f)|$ for $r = 0$; $-\int_{\mathbb{R}^d} f(x)\log f(x)dx$ for $r = 1$; and $-\log\|f\|_\infty$ for $r = \infty$. The case $r = 1$ corresponds to the classical Shannon differential entropy. Here, we denote by $|\mathrm{supp}(f)|$ the Lebesgue measure of the support of f, and $\|f\|_\infty$ represents the essential supremum of f. The r-Rényi entropy power is defined by

$$N_r(X) = e^{2h_r(X)/d}.$$

In the following, we drop the subscript when $r = 1$.

The classical Entropy Power Inequality (henceforth, EPI) of Shannon [39] and Stam [41], states that the entropy power $N(X)$ is super-additive on the sum of independent random vectors. There has been recent success in obtaining extensions of the EPI from the Shannon differential entropy to r-Rényi entropy. In [7, 8], Bobkov and Chistyakov showed that, at the expense of an absolute constant $c > 0$, the following Rényi EPI of order $r \in [1, \infty]$ holds

$$N_r(X_1 + \cdots + X_n) \geq c \sum_{i=1}^{n} N_r(X_i). \tag{4.2}$$

Ram and Sason soon after gave a sharpened constant depending on the number of summands [36]. Madiman, Melbourne, and Xu sharpened constants in the $r = \infty$ case by identifying extremizers in [31, 32]. Savaré and Toscani [38] showed that a modified Rényi entropy power is concave along the solution of a nonlinear heat equation, which generalizes Costa's concavity of entropy power [19]. Bobkov and Marsiglietti [10] proved the following variant of Rényi EPI

$$N_r(X + Y)^\alpha \geq N_r(X)^\alpha + N_r(Y)^\alpha \tag{4.3}$$

for $r > 1$ and some exponent α only depending on r. It is clear that (4.3) holds for more than two summands. Improvement of the exponent α was given by Li [27].

One of our goals is to establish analogues of (4.2) and (4.3) when the Rényi parameter $r \in (0, 1)$. Both (4.2) and (4.3) can be derived from Young's convolution inequality in conjunction with the entropic comparison inequality $h_{r_1}(X) \geq h_{r_2}(X)$ for any $0 \leq r_1 \leq r_2$. The latter fact is an immediate consequence of Jensen's inequality. When the Rényi parameter $r \in (0, 1)$, analogues of (4.2) and (4.3) require a converse of the entropic comparison inequality aforementioned. This technical issue prevents a general Rényi EPI of order $r \in (0, 1)$ for generic random vectors. Our first result shows that a general Rényi EPI of the form (4.2) indeed fails for all $r \in (0, 1)$.

Theorem 4.1 *For any $r \in (0, 1)$ and $\varepsilon > 0$, there exist independent random vectors $X_1, \cdots, X_n$ in $\mathbb{R}^d$, for some $d \geq 1$ and $n \geq 2$, such that*

$$N_r(X_1 + \cdots + X_n) < \varepsilon \sum_{i=1}^{n} N_r(X_i). \tag{4.4}$$

We have an explicit construction of such random vectors. They are essentially truncations of some spherically symmetric random vectors with finite covariance matrices and infinite Rényi entropies of order $r \in (0, 1)$. The key point is the convergence along the Central Limit Theorem (henceforth, CLT) for Rényi entropies of order $r \in (0, 1)$; that is, the r-Rényi entropy of their normalized sum converges to the r-Rényi entropy of a Gaussian. This implies that, after appropriate normalization, the LHS of (4.4) is finite, but the RHS of (4.4) can be as large as possible. The entropic CLT has been studied for a long time. A celebrated result of Barron [3] shows the convergence in the CLT for Shannon differential entropy (see [26] for a multidimensional setting). The recent work of Bobkov and Marsiglietti [11] studies the convergence in the CLT for Rényi entropy of order $r > 1$ for real-valued random variables (see also [12] for convergence in Rényi divergence, which is not equivalent to convergence in Rényi entropy unless $r = 1$). In Sect. 4.2, we establish the analogue of [11, Theorem 1.1] in higher dimensions and we prove convergence along the CLT for Rényi entropies of order $r \in (0, 1)$ for a large class of densities.

As mentioned above, the reverse entropic comparison inequality prevents Rényi EPIs of order $r \in (0, 1)$ for generic random vectors. However, a large class of random vectors with the so-called s-concave densities do satisfy such a reverse entropic comparison inequality. Our next results show that Rényi EPI of order $r \in (0, 1)$ holds for such densities. This extends the earlier work of Marsiglietti and Melbourne [33, 34] for log-concave densities (which corresponds to the $s = 0$ case).

Let $s \in [-\infty, \infty]$. A function $f : \mathbb{R}^d \to [0, \infty)$ is called s-concave if the inequality

$$f((1 - \lambda)x + \lambda y) \geq ((1 - \lambda) f(x)^s + \lambda f(y)^s)^{1/s} \tag{4.5}$$

holds for all $x, y \in \mathbb{R}^d$ such that $f(x)f(y) > 0$ and $\lambda \in (0, 1)$. For $s \in \{-\infty, 0, \infty\}$, the RHS of (4.5) is understood in the limiting sense; that is $\min\{f(x), f(y)\}$ for $s = -\infty$, $f(x)^{1-\lambda} f(y)^{\lambda}$ for $s = 0$, and $\max\{f(x), f(y)\}$ for $s = \infty$. The case $s = 0$ corresponds to log-concave functions. The study of measures with s-concave densities was initiated by Borell in the seminal work [13, 14]. One can think of s-concave densities, in particular log-concave densities, as functional versions of convex sets. There has been a recent stream of research on a formal parallel relation between functional inequalities of s-concave densities and geometric inequalities of convex sets.

Theorem 4.2 *For any $s \in (-1/d, 0)$ and $r \in (-sd, 1)$, there exists $c = c(s, r, d, n)$ such that for all independent random vectors $X_1, \cdots, X_n$ with s-*

concave densities in $\mathbb{R}^d$, *we have*

$$N_r(X_1 + \cdots + X_n) \geq c \sum_{i=1}^{n} N_r(X_i).$$

In particular, one can take

$$c = r^{\frac{1}{1-r}} \left(1 + \frac{1}{n|r'|}\right)^{1+n|r'|} \left(\prod_{k=1}^{d} \frac{(1+ks)^{|r'|(n-1)}(1+\frac{ks}{r})^{1+|r'|}}{(1+ks(1+\frac{1}{n|r'|}))^{1+n|r'|}}\right)^{\frac{2}{d}},$$

where $r' = r/(r-1)$ *is the Hölder conjugate of* r.

Theorem 4.3 *Given* $s \in (-1/d, 0)$, *there exist* $0 < r_0 < 1$ *and* $\alpha = \alpha(s, r, d)$ *such that for* $r \in (r_0, 1)$ *and independent random vectors* X *and* Y *in* $\mathbb{R}^d$ *with* s*-concave densities,*

$$N_r(X+Y)^\alpha \geq N_r(X)^\alpha + N_r(Y)^\alpha.$$

In particular, one can take

$$r_0 = \left(1 - \frac{2}{1+\sqrt{3}}\left(1 + \frac{1}{sd}\right)\right)^{-1}$$

$$\alpha = \left(1 + \frac{\log r + (r+1)\log\frac{r+1}{2r} + C(s)}{(1-r)\log 2}\right)^{-1},$$

where

$$C(s) = \frac{2}{d} \sum_{k=1}^{d} \left(\log\left(1 + \frac{ks}{r}\right) + r\log(1+ks) - (r+1)\log\left(1 + \frac{ks(r+1)}{2r}\right)\right).$$

Owing to the convexity, random vectors with s-concave densities also satisfy a reverse EPI, which was first proved by Bobkov and Madiman [9]. This can be seen as the functional lifting of Milman's well known reverse Brunn–Minkowski inequality [35]. Motivated by Busemann's theorem [17] in convex geometry, Ball et al. [2] conjectured that the following reverse EPI

$$N(X+Y)^{1/2} \leq N(X)^{1/2} + N(Y)^{1/2} \tag{4.6}$$

holds for any symmetric log-concave random vector $(X, Y) \in \mathbb{R}^2$. The r-Rényi entropy analogue was asked in [30], and the $r = 2$ case was soon verified in [27]. It was also observed in [27] that the r-Rényi entropy analogue is equivalent to the convexity of p-cross-section body in convex geometry introduced by Gardner and

Giannopoulos [23]. The equivalent linearization of (4.6) reads as follows. Let (X, Y) be a symmetric log-concave random vector in $\mathbb{R}^2$ such that $h(X) = h(Y)$. Then for any $\lambda \in [0, 1]$ we have

$$h((1-\lambda)X + \lambda Y) \leq h(X).$$

Cover and Zhang [20] proved the above inequality under the stronger assumption that X and Y have the same log-concave distribution. They also showed that this provides a characterization of log-concave distributions on the real line. The following theorem extends Cover and Zhang's result from log-concave densities to a more general class of s-concave densities. This gives an entropic characterization of s-concave densities and implies a reverse Rényi EPI for random vectors with the same s-concave density.

Theorem 4.4 *Let $r > 1 - 1/d$. Let f be a probability density function on $\mathbb{R}^d$. For any fixed integer $n \geq 2$, the identity*

$$\sup_{X_i \sim f} h_r\left(\sum_{i=1}^n \lambda_i X_i\right) = h_r(X_1)$$

holds for all $\lambda_i \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$ if and only if the density f is $(r-1)$-concave.

The paper is organized as follows. In Sect. 4.2, we explore the convergence along the CLT for r-Rényi entropies. For $r > 1$, the convergence is fully characterized for densities on $\mathbb{R}^d$, while for $r \in (0, 1)$ sufficient conditions are obtained for a large class of densities. More precisely, we prove the convergence for log-concave densities and for compactly supported, spherically symmetric and unimodal densities. As an application, we prove in Sect. 4.3 that a general r-Rényi EPI fails when $r \in (0, 1)$, thus establishing Theorem 4.1. We also complement this result by proving Theorems 4.2 and 4.3. In the last section, we provide an entropic characterization of the class of s-concave densities, and include a reverse Rényi EPI as an immediate consequence.

4.2 Convergence Along the CLT for Rényi Entropies

Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of independent identically distributed (henceforth, i.i.d.) centered random vectors in $\mathbb{R}^d$ with finite covariance matrix. We denote by Z_n the normalized sum

$$Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}. \tag{4.7}$$

An important tool used to prove various forms of CLT is the characteristic function. Recall that the characteristic function of a random vector X is defined by

$$\varphi_X(t) = \mathbb{E}\big[e^{i\langle t, X\rangle}\big], \quad t \in \mathbb{R}^d.$$

Before providing sufficient conditions for the convergence along the CLT for Rényi entropy of order $r \in (0, 1)$, we first extend [11, Theorem 1.1] to higher dimensions.

Theorem 4.5 *Let $r > 1$. Let $X_1, \cdots, X_n$ be i.i.d. centered random vectors in $\mathbb{R}^d$. We denote by ρ_n the density of Z_n defined in (4.7). The following statements are equivalent.*

1. *$h_r(Z_n) \to h_r(Z)$ as $n \to +\infty$, where Z is a Gaussian random vector with mean 0 and the same covariance matrix as X_1.*
2. *$h_r(Z_{n_0})$ is finite for some integer n_0.*
3. *$\int_{\mathbb{R}^d} |\varphi_{X_1}(t)|^\nu \, dt < +\infty$ for some $\nu \geq 1$.*
4. *Z_{n_0} has a bounded density ρ_{n_0} for some integer n_0.*

Proof $1 \Longrightarrow 2$: Assume that $h_r(Z_n) \to h_r(Z)$ as $n \to +\infty$. Then there exists an integer n_0 such that

$$h_r(Z) - 1 < h_r(Z_{n_0}) < h_r(Z) + 1.$$

Since $h_r(Z)$ is finite, we conclude that $h_r(Z_{n_0})$ is finite as well.

$2 \Longrightarrow 3$: Assume that $h_r(Z_{n_0})$ is finite for some integer n_0. Then Z_{n_0} has a density $\rho_{n_0} \in L^r(\mathbb{R}^d)$.

Case 1 If $r \geq 2$, we have $\rho_{n_0} \in L^2(\mathbb{R}^d)$. Using Plancherel's identity, we have $\varphi_{Z_{n_0}} \in L^2(\mathbb{R}^d)$. It follows that

$$\int_{\mathbb{R}^d} |\varphi_{Z_{n_0}}(t)|^2 \, dt = \int_{\mathbb{R}^d} |\varphi_{X_1}\left(t/\sqrt{n_0}\right)|^{2n_0} \, dt < +\infty.$$

For $\nu = 2n_0$, we have

$$\int_{\mathbb{R}^d} |\varphi_{X_1}(t)|^\nu \, dt < +\infty.$$

Case 2 If $r \in (1, 2)$, we apply the Hausdorff–Young inequality to obtain

$$\|\varphi_{Z_{n_0}}\|_{L^{r'}} \leq \frac{1}{(2\pi)^{d/r'}} \|\rho_{n_0}\|_{L^r},$$

where r' is the conjugate of r such that $1/r + 1/r' = 1$. Hence, for $\nu = r'n_0$, we have

$$\int_{\mathbb{R}^d} |\varphi_{X_1}(t)|^\nu \, dt < +\infty.$$

3 $\Longrightarrow$ 4: Since $\int_{\mathbb{R}^d} |\varphi_{X_1}(t)|^\nu \, dt < +\infty$ for some $\nu \geq 1$, one may apply Gnedenko's local limit theorems (see [24]), which is valid in arbitrary dimensions (see [5]). In particular, we have

$$\lim_{n\to+\infty} \sup_{x\in\mathbb{R}^d} |\rho_n(x) - \phi_\Sigma(x)| = 0, \tag{4.8}$$

where ϕ_Σ denotes the density of a Gaussian random vector with mean 0 and the same covariance matrix as X_1. We deduce that there exists an integer n_0 and a constant $M > 0$ such that $\rho_n \leq M$ for all $n \geq n_0$.

4 $\Longrightarrow$ 1: Since ρ_{n_0} is bounded, then $\rho_{n_0} \in L^2$, and we deduce by Plancherel's identity that $\int_{\mathbb{R}^d} |\varphi_{X_1}(t)|^\nu \, dt < +\infty$ for $\nu = 2n_0$. Hence, (4.8) holds and there exists $M > 0$ such that $\rho_n \leq M$ for all $n \geq n_0$. Let us show that $\int_{\mathbb{R}^d} \rho_n(x)^r dx \to \int_{\mathbb{R}^d} \phi_\Sigma(x)^r dx$ as $n \to +\infty$, where ϕ_Σ denotes the density of a Gaussian random vector with mean 0 and the same covariance matrix as X_1. By the CLT, for any $\varepsilon > 0$, there exists $T > 0$ such that for all n large enough,

$$\int_{|x|>T} \rho_n(x) dx < \varepsilon,$$

which implies that

$$\int_{|x|>T} \rho_n(x)^r dx \leq M^{r-1} \int_{|x|>T} \rho_n(x) dx < M^{r-1}\varepsilon.$$

The function ϕ_Σ satisfies similar inequalities. Hence, for any $\delta > 0$, there exists $T > 0$ such that for all n large enough,

$$\left| \int_{|x|>T} \rho_n(x)^r dx - \int_{|x|>T} \phi_\Sigma(x)^r dx \right| < \delta.$$

On the other hand, by (4.8), for all $T > 0$, the function $\rho_n^r(x)\mathbf{1}_{\{|x|\leq T\}}$ converges everywhere to $\phi_\Sigma^r(x)\mathbf{1}_{\{|x|\leq T\}}$ as $n \to +\infty$. Since $\rho_n^r(x)\mathbf{1}_{\{|x|\leq T\}}$ is dominated by the integrable function $M^r\mathbf{1}_{\{|x|\leq T\}}$, one may use the Lebesgue dominated theorem to conclude that

$$\lim_{n\to+\infty} \left| \int_{|x|\leq T} \rho_n(x)^r dx - \int_{|x|\leq T} \phi_\Sigma(x)^r dx \right| = 0.$$

□

Remark 4.6 Theorem 4.5 fails for $r \in (0, 1)$. For example, one can consider i.i.d. random vectors with a bounded density $\rho(x)$ such that $\int_{\mathbb{R}^d} \rho(x)^r dx = +\infty$ (e.g., Cauchy-type distributions). The implication 4 $\Longrightarrow$ 2 (and thus 4 $\Longrightarrow$ 1) will not hold since by Jensen inequality $h_r(Z_n) \geq h_r(X_1/\sqrt{n}) = \infty$ for all $n \geq 1$. As observed by Barron [3], the implication 1 $\Longrightarrow$ 4 does not necessarily hold in the Shannon entropy case $r = 1$.

The following result yields a sufficient condition for convergence along the CLT to hold for Rényi entropies of order $r \in (0, 1)$ for a large class of random vectors in $\mathbb{R}^d$.

Theorem 4.7 *Let $r \in (0, 1)$. Let $X_1, \cdots, X_n$ be i.i.d. centered log-concave random vectors in $\mathbb{R}^d$. Then we have $h_r(Z_n) < +\infty$ for all $n \geq 1$, and*

$$\lim_{n\to\infty} h_r(Z_n) = h_r(Z),$$

where Z_n is the normalized sum in (4.7) and Z is a Gaussian random vector with mean 0 and the same covariance matrix as X_1.

Proof Since log-concavity is preserved under independent sum, Z_n is log-concave for all $n \geq 1$. Hence, for all $n \geq 1$, Z_n has a bounded log-concave density ρ_n, which satisfies

$$\rho_n(x) \leq e^{-a_n|x|+b_n},$$

for all $x \in \mathbb{R}^d$, and for some constants $a_n > 0$, $b_n \in \mathbb{R}$ possibly depending on the dimension (see, e.g., [16]). Hence, for all $n \geq 1$, we have

$$\int_{\mathbb{R}^d} \rho_n(x)^r \, dx \leq \int_{\mathbb{R}^d} e^{-r(a_n|x|+b_n)} \, dx < +\infty.$$

We deduce that $h_r(Z_n) < +\infty$ for all $n \geq 1$.

The boundedness of ρ_n implies that (4.8) holds, and thus there exists an integer n_0 such that for all $n \geq n_0$,

$$\rho_n(0) > \frac{1}{2}\phi_\Sigma(0),$$

where Σ is the covariance matrix of X_1 (and thus does not depend on n). Moreover, since ρ_n is log-concave, one has for all $x \in \mathbb{R}^d$ that

$$\rho_n(rx) = \rho_n((1-r)0 + rx) \geq \rho_n(0)^{1-r}\rho_n(x)^r \geq \frac{1}{2^{1-r}}\phi_\Sigma(0)^{1-r}\rho_n(x)^r.$$

Hence, for all $T > 0$, we have

$$\begin{aligned}\int_{|x|>T} \rho_n(x)^r \, dx &\leq \frac{2^{1-r}}{\phi_\Sigma(0)^{1-r}} \int_{|x|>T} \rho_n(rx) \, dx \\ &= \frac{2^{1-r}}{r^d \phi_\Sigma(0)^{1-r}} \mathbb{P}(|Z_n| > rT) \\ &\leq \frac{1}{T^2} \frac{2^{1-r}\mathbb{E}[|X_1|^2]}{r^{d+2}\phi_\Sigma(0)^{1-r}},\end{aligned}$$

where the last inequality follows from Markov's inequality and the fact that

$$\mathbb{E}[|Z_n|^2] = \frac{\mathbb{E}[|X_1|^2] + \cdots + \mathbb{E}[|X_n|^2]}{n} = \mathbb{E}[|X_1|^2].$$

Hence, for every $\varepsilon > 0$, one may choose a positive number T such that for all n large enough,

$$\int_{|x|>T} \rho_n(x)^r dx < \varepsilon, \qquad \int_{|x|>T} \phi_\Sigma(x)^r dx < \varepsilon,$$

and hence

$$\left| \int_{|x|>T} \rho_n(x)^r dx - \int_{|x|>T} \phi_\Sigma(x)^r dx \right| < \varepsilon.$$

On the other hand, from (4.8), we conclude as in the proof of Theorem 4.5 that for all $T > 0$,

$$\lim_{n\to+\infty} \left| \int_{|x|\le T} \rho_n(x)^r dx - \int_{|x|\le T} \phi_\Sigma(x)^r dx \right| = 0.$$

□

A function $f : \mathbb{R}^d \to \mathbb{R}$ is called unimodal if the super-level sets $\{x \in \mathbb{R}^d : f(x) > t\}$ are convex for all $t \in \mathbb{R}$. Next, we provide a convergence result for random vectors in $\mathbb{R}^d$ with unimodal densities under additional symmetry assumptions. First, we need the following stability result.

Proposition 4.8 *The class of spherically symmetric and unimodal random variables is stable under convolution.*

Proof Let f_1 and f_2 be two spherically symmetric and unimodal densities. By assumption, f_i satisfy that $f_i(Tx) = f_i(x)$ for an orthogonal map T and $|x| \le |y|$ implies $f_i(x) \ge f_i(y)$. By the layer cake decomposition, we write

$$f_i(x) = \int_0^\infty \mathbf{1}_{\{(u,v):f_i(u)>v\}}(x, \lambda) d\lambda.$$

Apply Fubini's theorem to obtain

$$\begin{aligned} f_1 \star f_2(x) &= \int_{\mathbb{R}^d} f_1(x-y) f_2(y) dy \\ &= \int_0^\infty \int_0^\infty \left(\int_{\mathbb{R}^d} \mathbf{1}_{\{(u,v):f_1(u)>v\}}(x-y, \lambda_1) \mathbf{1}_{\{(u,v):f_2(u)>v\}}(y, \lambda_2) dy \right) \\ &\quad \times d\lambda_1 d\lambda_2. \end{aligned} \tag{4.9}$$

Notice that by the spherical symmetry and decreasingness of f_i, the super-level set

$$L_{\lambda_i} = \{u : f_i(u) > \lambda_i\}$$

is an origin symmetric ball. Thus we can write the integrand in (4.9) as

$$\int_{\mathbb{R}^d} \mathbf{1}_{L_{\lambda_1}}(x-y)\mathbf{1}_{L_{\lambda_2}}(y)dy = \mathbf{1}_{L_{\lambda_1}} \star \mathbf{1}_{L_{\lambda_2}}(x).$$

This quantity is clearly dependent only on $|x|$, giving spherical symmetry. In addition, as the convolution of two log-concave functions, $\mathbf{1}_{L_{\lambda_1}} \star \mathbf{1}_{L_{\lambda_2}}$ is log-concave as well. It follows that for every λ_1, λ_2, and $|x| \leq |y|$ we have

$$\mathbf{1}_{L_{\lambda_1}} \star \mathbf{1}_{L_{\lambda_2}}(x) \geq \mathbf{1}_{L_{\lambda_1}} \star \mathbf{1}_{L_{\lambda_2}}(y).$$

Integrating this inequality completes the proof. □

Let us establish large deviation and pointwise inequalities for compactly supported, spherically symmetric and unimodal densities.

Theorem 4.9 (Hoeffding [25]) *Let $X_1, \cdots, X_n$ be independent random variables with mean 0 and bounded in (a_i, b_i), respectively. One has for all $T > 0$,*

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i > T\right) \leq \exp\left(-\frac{2T^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

The following result is Hoeffding's inequality in higher dimensions.

Lemma 4.10 *Let $X_1, \cdots, X_n$ be centered independent random vectors in $\mathbb{R}^d$ satisfying $\mathbb{P}(|X_i| > R) = 0$ for some $R > 0$. One has for all $T > 0$ that*

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right| > T\right) \leq 2d \exp\left(-\frac{T^2}{2d^2R^2}\right).$$

Proof Let $X_{i,j}$ be the j-th coordinate of the random vector X_i. Then we have

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right| > T\right) \leq \mathbb{P}\left(\bigcup_{j=1}^{d}\left\{|X_{1,j} + \cdots + X_{n,j}| > \frac{T\sqrt{n}}{d}\right\}\right) \tag{4.10}$$

$$\leq \sum_{j=1}^{d} \mathbb{P}\left(|X_{1,j} + \cdots + X_{n,j}| > \frac{T\sqrt{n}}{d}\right) \tag{4.11}$$

$$\leq 2d \exp\left(-\frac{T^2}{2d^2R^2}\right), \tag{4.12}$$

where inequality (4.10) follows from the pigeon-hole principle, (4.11) from a union bound, and (4.12) follows from applying Theorem 4.9 to $X_{1,j} + \cdots + X_{n,j}$ and $(-X_{1,j}) + \cdots + (-X_{n,j})$. □

We deduce the following pointwise estimate for unimodal spherically symmetric and bounded random variables.

Corollary 4.11 *Let $X_1, \cdots, X_n$ be i.i.d. random vectors with spherically symmetric, unimodal density supported on the Euclidean ball $B_R = \{x : |x| \le R\}$ for some $R > 0$. Let ρ_n denote the density of the normalized sum Z_n. Then there exists $c_d > 0$ such that for all $n \ge 1$ and $|x| > 2$,*

$$\rho_n(x) \le c_d \exp\left(-\frac{(|x|-1)^2}{2d^2R^2}\right).$$

Proof Stating Lemma 4.10 in terms of ρ_n, we have

$$\int_{|w|>T} \rho_n(w)dw \le 2d \exp\left(-\frac{T^2}{2d^2R^2}\right). \tag{4.13}$$

Since the class of spherically symmetric unimodal random variables is stable under independent summation by Proposition 4.8, ρ_n is spherically symmetric and unimodal, so that

$$\begin{aligned}\rho_n(x) &\le \frac{\int_{B_{|x|}\setminus B_{|x|-1}} \rho_n(w)dw}{\mathrm{Vol}(B_{|x|}\setminus B_{|x|-1})} \\ &\le \frac{\int_{|w|\ge|x|-1} \rho_n(w)dw}{(2^d-1)\omega_d} \end{aligned}\tag{4.14}$$

where $B_{|x|}$ represents the Euclidean ball of radius $|x|$ centered at the origin and ω_d is the volume of the unit ball. Note that

$$\mathrm{Vol}(B_{|x|}\setminus B_{|x|-1}) = (|x|^d - (|x|-1)^d)\omega_d \ge (2^d-1)\omega_d,$$

since $t \mapsto t^d - (t-1)^d$ is increasing, so that (4.14) follows. Now applying (4.13) we have

$$\begin{aligned}\rho_n(x) &\le \frac{\int_{|w|\ge|x|-1} \rho_n(w)dw}{(2^d-1)\omega_d} \\ &\le \frac{2d}{(2^d-1)\omega_d} \exp\left(-\frac{(|x|-1)^2}{2d^2R^2}\right)\end{aligned}$$

and our result holds with

$$c_d = \frac{2d}{(2^d - 1)\omega_d}.$$

□

We are now ready to establish a convergence result for bounded spherically symmetric unimodal random vectors.

Theorem 4.12 *Let $r \in (0, 1)$. Let $X_1, \cdots, X_n$ be i.i.d. random vectors in $\mathbb{R}^d$ with a spherically symmetric unimodal density with compact support. Then we have*

$$\lim_{n\to\infty} h_r(Z_n) = h_r(Z),$$

where Z_n is the normalized sum in (4.7) and Z is a Gaussian random vector with mean 0 and the same covariance matrix as X_1.

Proof Let us denote by ρ_n the density of Z_n. Since ρ_1 is bounded, one may apply (4.8) together with Lebesgue dominated convergence to conclude that for all $T > 0$,

$$\lim_{n\to+\infty} \left| \int_{|x|\le T} \rho_n(x)^r dx - \int_{|x|\le T} \phi_\Sigma(x)^r dx \right| = 0.$$

On the other hand, by Corollary 4.11, one may choose $T > 0$ such that for all $n \ge 1$,

$$\int_{|x|>T} \rho_n(x)^r dx < \varepsilon, \qquad \int_{|x|>T} \phi_\Sigma(x)^r dx < \varepsilon,$$

and hence

$$\left| \int_{|x|>T} \rho_n(x)^r dx - \int_{|x|>T} \phi_\Sigma(x)^r dx \right| < \varepsilon.$$

□

4.3 Rényi EPIs of Order $r \in (0, 1)$

A striking difference between Rényi EPIs of orders $r \in (0, 1)$ and $r \ge 1$ is the lack of an absolute constant. Indeed, it was shown in [8] that for $r \ge 1$ Rényi EPI of the form (4.2) holds for generic independent random vectors with an absolute constant $c \ge \frac{1}{e} r^{\frac{1}{r-1}}$. In the following subsection, we show that such a Rényi EPI does not hold for $r \in (0, 1)$.

4.3.1 Failure of a Generic Rényi EPI

Definition 4.13 For $r \in [0, \infty]$, we define c_r as the largest number such that for all $n, d \geq 1$ and any independent random vectors $X_1, \cdots, X_n$ in $\mathbb{R}^d$, we have

$$N_r(X_1 + \cdots + X_n) \geq c_r \sum_{i=1}^{n} N_r(X_i). \tag{4.15}$$

Then we can rephrase Theorem 4.1 as follows.

Theorem 4.14 *For $r \in (0, 1)$, the constant c_r defined in (4.15) satisfies $c_r = 0$.*

The motivating observation for this line of argument is the fact that for $r \in (0, 1)$, there exist distributions with finite covariance matrices and infinite r-Rényi entropies. One might anticipate that this could contradict the existence of an r-Rényi EPI, as the CLT forces the normalized sum of i.i.d. random vectors $X_1, \cdots, X_n$ drawn from such a distribution to become "more Gaussian". Heuristically, one anticipates that $N_r(X_1 + \cdots + X_n)/n = N_r(Z_n)$ should approach $N_r(Z)$ for large n, where Z_n is the normalized sum in (4.7) and Z is a Gaussian vector with the same covariance matrix as X_1, while $\sum_{i=1}^{n} N_r(X_i)/n = N_r(X_1)$ is infinite.

Proof of Theorem 4.14 Let us consider the following density

$$f_{R,p,d}(x) = C_R(1 + |x|)^{-p}\mathbf{1}_{B_R}(x) \quad x \in \mathbb{R}^d,$$

with $p, R > 0$ and C_R implicitly determined to make $f_{R,p,d}$ a density. Since the density is spherically symmetric, its covariance matrix can be rewritten as $\sigma_R^2 I$ for some $\sigma_R > 0$, where I is the identity matrix. Computing in spherical coordinates one can check that $\lim_{R\to\infty} C_R$ is finite for $p > d$, and we can thus define a density $f_{\infty,p,d}$. What is more, when $p > d + 2$, the limiting density $f_{\infty,p,d}$ has a finite covariance matrix, and has finite Rényi entropy if and only if $p > d/r$.

For fixed $r \in (0, 1)$, we take $p \in (d^* + 2, d^*/r]$, where $d^* = \min\{d \in \mathbb{N} : d > 2r/(1 - r)\}$ guarantees the existence of such p. In this case, the limiting density f_{∞,p,d^*} is well defined and it has finite covariance matrix $\sigma_\infty^2 I$, but the corresponding r-Rényi entropy is infinite. Now we select independent random vectors $X_1, \cdots, X_n$ from the distribution f_{R,p,d^*}. Since f_{R,p,d^*} is a spherically symmetric and unimodal density with compact support, we can apply Theorem 4.12 to conclude that

$$\lim_{n\to\infty} N_r(Z_n) = \sigma_R^2 N_r(Z_{Id}),$$

where Z_n is the normalized sum in (4.7) and Z_{Id} is the standard d-dimensional Gaussian. Since $\lim_{R\to\infty} \sigma_R = \sigma_\infty < \infty$, we can take R large enough such that

$|\sigma_R^2 - \sigma_\infty^2| \leq 1$. Then we can take n large enough such that

$$N_r(Z_n) \leq (\sigma_\infty^2 + 2) N_r(Z_{Id}). \tag{4.16}$$

Since the limiting density f_{∞,p,d^*} has infinite r-Rényi entropy, given $M > 0$, we can take R large enough such that

$$N_r(X_1) \geq M. \tag{4.17}$$

Combining (4.16) and (4.17), we conclude that for inequality (4.15) to hold we must have

$$c_r \leq \frac{(\sigma_\infty^2 + 2) N_r(Z_{Id})}{M}$$

for all $M > 0$. Then the statement follows from taking the limit $M \to \infty$. □

Remark 4.15 Random vectors in our proof has identical s-concave density with $s \leq -r/d$. In the following section, we provide a complementary result by showing that Rényi EPI of order $r \in (0, 1)$ does hold for s-concave densities when $-r/d < s < 0$.

4.3.2 Rényi EPIs for s-Concave Densities

As showed above, a generic Rényi EPI of the form (4.2) fails for $r \in (0, 1)$. In this part, we establish Rényi EPIs of the forms (4.2) and (4.3) for an important class of random vectors with s-concave densities (see (4.5)).

Following Lieb [29], we prove Theorems 4.2 and 4.3 by showing their equivalent linearizations. The following linearization of (4.2) and (4.3) is due to Rioul [37]. The $c = 1$ case was used in [27].

Theorem 4.16 ([37]) *Let $X_1, \cdots, X_n$ be independent random vectors in $\mathbb{R}^d$. The following statements are equivalent.*

1. *There exist a constant $c > 0$ and an exponent $\alpha > 0$ such that*

$$N_r^\alpha\left(\sum_{i=1}^n X_i\right) \geq c \sum_{i=1}^n N_r^\alpha(X_i). \tag{4.18}$$

2. *For any $\lambda_1, \cdots, \lambda_n \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$, one has*

$$h_r\left(\sum_{i=1}^n \sqrt{\lambda_i} X_i\right) - \sum_{i=1}^n \lambda_i h_r(X_i) \geq \frac{d}{2}\left(\frac{\log c}{\alpha} + \left(\frac{1}{\alpha} - 1\right) H(\lambda)\right), \tag{4.19}$$

where $H(\lambda) \triangleq H(\lambda_1, \cdots, \lambda_n)$ *is the discrete entropy defined as*

$$H(\lambda) = -\sum_{i=1}^{n} \lambda_i \log \lambda_i .$$

Inequality (4.19) is the linearized form of inequality (4.18). One of the ingredients used to establish (4.19) is Young's sharp convolution inequality [4, 15]. Its information-theoretic formulation was given in [21], which we recall below. We denote by r' the Hölder conjugate of r such that $1/r + 1/r' = 1$.

Theorem 4.17 ([15, 21]) *Let* $r > 0$. *Let* $\lambda_1, \cdots, \lambda_n \geq 0$ *such that* $\sum_{i=1}^{n} \lambda_i = 1$, *and let* $r_1, \cdots, r_n$ *be positive reals such that* $\lambda_i = r'/r_i'$. *For any independent random vectors* $X_1, \cdots, X_n$ *in* $\mathbb{R}^d$, *one has*

$$h_r\left(\sum_{i=1}^{n} \sqrt{\lambda_i} X_i\right) - \sum_{i=1}^{n} \lambda_i h_{r_i}(X_i) \geq \frac{d}{2} r' \left(\frac{\log r}{r} - \sum_{i=1}^{n} \frac{\log r_i}{r_i}\right). \tag{4.20}$$

The second ingredient is a comparison between Rényi entropies h_r and h_{r_i}. When $r > 1$, we have $1 < r_i < r$, and Jensen's inequality implies that $h_r \leq h_{r_i}$. In this case, one can deduce (4.19) from (4.20) with h_{r_i} replaced by h_r. However, when $r \in (0, 1)$, the order of r and r_i are reversed, i.e., $0 < r < r_i < 1$, and we need a reverse entropy comparison inequality. The so-called s-concave densities do satisfy such a reverse entropy comparison inequality. The following result of Fradelizi et al. [22] serves this purpose.

Theorem 4.18 ([22]) *Let* $s \in \mathbb{R}$. *Let* $f : \mathbb{R}^d \to [0, +\infty)$ *be an integrable* s-*concave function. The function*

$$G(r) = C(r) \int_{\mathbb{R}^d} f(x)^r \, dx$$

is log-concave for $r > \max\{0, -sd\}$, *where*

$$C(r) = (r + s) \cdots (r + sd). \tag{4.21}$$

We deduce the following Rényi entropic comparison for random vectors with s-concave densities.

Corollary 4.19 *Let* X *be a random vector in* $\mathbb{R}^d$ *with a* s-*concave density. For* $-sd < r < q < 1$, *we have*

$$h_q(X) \geq h_r(X) + \log \frac{C(r)^{\frac{1}{1-r}} C(1)^{\frac{q-r}{(1-q)(1-r)}}}{C(q)^{\frac{1}{1-q}}}.$$

Proof Write $q = (1-\lambda)\cdot r + \lambda \cdot 1$. Using the log-concavity of the function G in Theorem 4.18, we have

$$G(q) \geq G(r)^{1-\lambda} G(1)^{\lambda} = G(r)^{\frac{1-q}{1-r}} G(1)^{\frac{q-r}{1-r}}.$$

The above inequality can be rewritten in terms of entropy power as follows

$$C(q)^{\frac{2}{d}\cdot\frac{1}{1-q}} N_q(X) \geq C(r)^{\frac{2}{d}\cdot\frac{1-q}{1-r}\cdot\frac{1}{1-q}} N_r(X) C(1)^{\frac{2}{d}\cdot\frac{q-r}{1-r}\cdot\frac{1}{1-q}}.$$

The desired statement follows from taking the logarithm of both sides of the above inequality. □

Theorem 4.17 together with Corollary 4.19 yields the following Rényi EPI with a single Rényi parameter $r \in (0,1)$ for s-concave densities.

Theorem 4.20 *Let $s \in (-1/d, 0)$ and $r \in (-sd, 1)$. Let $X_1, \cdots, X_n$ be independent random vectors in $\mathbb{R}^d$ with s-concave densities. For all $\lambda = (\lambda_1, \cdots, \lambda_n) \in [0,1]^n$ such that $\sum_{i=1}^n \lambda_i = 1$, we have*

$$h_r\left(\sum_{i=1}^n \sqrt{\lambda_i} X_i\right) - \sum_{i=1}^n \lambda_i h_r(X_i) \geq \frac{d}{2} A(\lambda) + \sum_{k=1}^d g_k(\lambda),$$

where

$$A(\lambda) = r'\left(\left(1-\frac{1}{r'}\right)\log\left(1-\frac{1}{r'}\right) - \sum_{i=1}^n \left(1-\frac{\lambda_i}{r'}\right)\log\left(1-\frac{\lambda_i}{r'}\right)\right),$$

$$g_k(\lambda) = (1-n)r'\log(1+ks) + (1-r')\log\left(1+\frac{ks}{r}\right) + r'\sum_{i=1}^n \left(1-\frac{\lambda_i}{r'}\right) \times \log\left(1+ks\left(1-\frac{\lambda_i}{r'}\right)\right).$$

Proof Let r_i be defined by $\lambda_i = r'/r_i'$, where r' and r_i' are Hölder conjugates of r and r_i, respectively. Combining Theorem 4.17 with Corollary 4.19, we have

$$h_r\left(\sum_{i=1}^n \sqrt{\lambda_i} X_i\right) - \sum_{i=1}^n \lambda_i h_r(X_i) \geq \frac{d}{2} r'\left(\frac{\log r}{r} - \sum_{i=1}^n \frac{\log r_i}{r_i}\right) + \sum_{i=1}^n \lambda_i \log \frac{C(r)^{\frac{1}{1-r}} C(1)^{\frac{r_i - r}{(1-r_i)(1-r)}}}{C(r_i)^{\frac{1}{1-r_i}}}. \tag{4.22}$$

Notice that $C(r) = r^d D(r)$, where $C(r)$ is given in (4.21) and $D(r) = (1 + s/r) \cdots (1 + sd/r)$. Thus,

$$\sum_{i=1}^{n} \lambda_i \log \frac{C(r)^{\frac{1}{1-r}} C(1)^{\frac{r_i - r}{(1-r_i)(1-r)}}}{C(r_i)^{\frac{1}{1-r_i}}}$$
$$= \sum_{i=1}^{n} \lambda_i \left(\frac{\log D(r)}{1-r} + \left(\frac{1}{1-r_i} - \frac{1}{1-r} \right) \log D(1) - \frac{\log D(r_i)}{1-r_i} \right)$$
$$+ d \left(\frac{\log r}{1-r} - \sum_{i=1}^{n} \lambda_i \frac{\log r_i}{1-r_i} \right). \tag{4.23}$$

Using the identities $1/(1-r) = 1 - r'$ and $\lambda_i/(1-r_i) = \lambda_i - r'$, we have

$$\sum_{i=1}^{n} \lambda_i \left(\frac{\log D(r)}{1-r} + \left(\frac{1}{1-r_i} - \frac{1}{1-r} \right) \log D(1) - \frac{\log D(r_i)}{1-r_i} \right)$$
$$= (1-r') \log D(r) + (1-n) r' \log D(1) + \sum_{k=1}^{d} \sum_{i=1}^{n} (r' - \lambda_i) \log \left(1 + \frac{ks}{r_i} \right)$$
$$= \sum_{k=1}^{d} \left((1-r') \log \left(1 + \frac{ks}{r} \right) + (1-n) r' \log(1+ks) \right.$$
$$\left. + \sum_{i=1}^{n} (r' - \lambda_i) \log \left(1 + \frac{ks}{r_i} \right) \right) = \sum_{k=1}^{d} g_k(\lambda). \tag{4.24}$$

The last identity follows from $1/r_i = 1 - \lambda_i/r'$. Using (4.24) and (4.23), the RHS of (4.22) can be written as

$$\frac{d}{2} r' \left(\frac{\log r}{r} - \sum_{i=1}^{n} \frac{\log r_i}{r_i} \right) + d \left(\frac{\log r}{1-r} - \sum_{i=1}^{n} \lambda_i \frac{\log r_i}{1-r_i} \right)$$
$$+ \sum_{k=1}^{d} g_k(\lambda) = \frac{d}{2} A(\lambda) + \sum_{k=1}^{d} g_k(\lambda).$$

This concludes the proof. □

Having Theorems 4.16 and 4.20 at hand, we are ready to prove Theorems 4.2 and 4.3.

4.3.2.1 Proof of Theorem 4.2

Put Theorems 4.16 and 4.20 together. Then it suffices to find c such that the following inequality

$$\frac{d}{2}A(\lambda) + \sum_{k=1}^{d} g_k(\lambda) \geq \frac{d}{2}\log c$$

holds for all $\lambda = (\lambda_1, \cdots, \lambda_n) \in [0, 1]^n$ such that $\sum_{i=1}^{n} \lambda_i = 1$. Hence, we can set

$$c = \inf_{\lambda} \exp\left(A(\lambda) + \frac{2}{d}\sum_{k=1}^{d} g_k(\lambda)\right),$$

where the infimum runs over all $\lambda = (\lambda_1, \cdots, \lambda_n) \in [0, 1]^n$ such that $\sum_{i=1}^{n} \lambda_i = 1$. For fixed r, both $A(\lambda)$ and $g_k(\lambda)$ are sum of one-dimensional convex functions of the form $(1 + x)\log(1 + x)$. Furthermore, both $A(\lambda)$ and $g_k(\lambda)$ are permutation invariant. Hence, the minimum is achieved at $\lambda = (1/n, \cdots, 1/n)$. This yields the numerical value of c in Theorem 4.2.

4.3.2.2 Proof of Theorem 4.3

The following lemma in [33] serves us in the proof of Theorem 4.3.

Lemma 4.21 ([33]) *Let* $c > 0$. *Let* $L, F : [0, c] \to [0, \infty)$ *be twice differentiable on* $(0, c]$, *continuous on* $[0, c]$, *such that* $L(0) = F(0) = 0$ *and* $L'(c) = F'(c) = 0$. *Let us also assume that* $F(x) > 0$ *for* $x > 0$, *that* F *is strictly increasing, and that* F' *is strictly decreasing. Then* $\frac{L''}{F''}$ *increasing on* $(0, c)$ *implies that* $\frac{L}{F}$ *is increasing on* $(0, c)$ *as well. In particular,*

$$\max_{x\in[0,c]} \frac{L(x)}{F(x)} = \frac{L(c)}{F(c)}.$$

Proof of Theorem 4.3 Apply Theorems 4.16 and 4.20 with $n = 2$. Then it suffices to find α such that for all $\lambda \in [0, 1]$ we have

$$\frac{d}{2}A(\lambda) + \sum_{k=1}^{d} g_k(\lambda) \geq \frac{d}{2}\left(\frac{1}{\alpha} - 1\right)H(\lambda),$$

where

$$A(\lambda) = r'\left(\left(1-\frac{1}{r'}\right)\log\left(1-\frac{1}{r'}\right)-\left(1-\frac{\lambda}{r'}\right)\log\left(1-\frac{\lambda}{r'}\right)-\left(1-\frac{1-\lambda}{r'}\right)\right.$$
$$\left.\times\log\left(1-\frac{1-\lambda}{r'}\right)\right),$$
$$g_k(\lambda) = (1-r')\log\left(1+\frac{ks}{r}\right)-r'\log(1+ks)$$
$$+r'\left(\left(1-\frac{\lambda}{r'}\right)\log\left(1+ks\left(1-\frac{\lambda}{r'}\right)\right)+\left(1-\frac{1-\lambda}{r'}\right)\right.$$
$$\left.\times\log\left(1+ks\left(1-\frac{1-\lambda}{r'}\right)\right)\right).$$

We can set

$$\alpha = \left(1-\sup_{0\le\lambda\le 1}\left(-\frac{A(\lambda)}{H(\lambda)}-\frac{2}{d}\sum_{k=1}^{d}\frac{g_k(\lambda)}{H(\lambda)}\right)\right)^{-1}. \tag{4.25}$$

We will show that the optimal value is achieved at $\lambda = 1/2$. Since the function is symmetric about $\lambda = 1/2$, it suffices to show that

$$-\frac{A(\lambda)}{H(\lambda)}-\frac{2}{d}\sum_{k=1}^{n}\frac{g_k(\lambda)}{H(\lambda)} \tag{4.26}$$

is increasing on $[0, 1/2]$. It has been shown in [27] that $-A(\lambda)/H(\lambda)$ is increasing on $[0, 1/2]$. We will show that for each $k = 1, \cdots, n$ the function $-g_k(\lambda)/H(\lambda)$ is also increasing on $[0, 1/2]$. One can check that $-g_k(\lambda)$ and $H(\lambda)$ satisfy the conditions in Lemma 4.21. Hence, it suffices to show that $-g_k''(\lambda)/H''(\lambda)$ is increasing on $[0, 1/2]$. Elementary calculation yields that

$$H''(\lambda) = -\frac{1}{\lambda(1-\lambda)}.$$

Define $x = \frac{\lambda}{|r'|}$ and $y = \frac{1-\lambda}{|r'|} = \frac{1}{|r'|} - x$. Then one can check that

$$-g_k''(\lambda) = \frac{ks}{|r'|}\left(\frac{1}{1+ks(1+x)}+\frac{1}{1+ks(1+y)}+\frac{1}{(1+ks(1+x))^2}\right.$$
$$\left.+\frac{1}{(1+ks(1+y))^2}\right).$$

Hence, we have

$$-\frac{g_k''(\lambda)}{H''(\lambda)} = ksr'W(x),$$

where

$$W(x) = xy\left(\frac{1}{1+ks(1+x)} + \frac{1}{1+ks(1+y)} + \frac{1}{(1+ks(1+x))^2} + \frac{1}{(1+ks(1+y))^2}\right).$$

Since $s, r' < 0$, it suffices to show that $W(x)$ is increasing on $[0, \frac{1}{2|r'|}]$. We rewrite W as follows

$$W(x) = W_1(x) + W_2(x),$$

where

$$W_1(x) = xy\left(\frac{1}{1+ks(1+x)} + \frac{1}{1+ks(1+y)}\right),$$

$$W_2(x) = xy\left(\frac{1}{(1+ks(1+x))^2} + \frac{1}{(1+ks(1+y))^2}\right). \tag{4.27}$$

We will show that both $W_1(x)$ and $W_2(x)$ are increasing on $[0, \frac{1}{2|r'|}]$.

Now let us focus on W_1. Since $y = \frac{1}{|r'|} - x$, one can check that

$$W_1'(x) = \left(\frac{1}{|r'|} - 2x\right)\left(\frac{1}{1+ks(1+x)} + \frac{1}{1+ks(1+y)}\right) - ksxy\left(\frac{1}{(1+ks(1+x))^2} - \frac{1}{(1+ks(1+y))^2}\right).$$

Let us denote

$$a \triangleq a(x) = 1 + ks(1+x), \tag{4.28}$$

$$b \triangleq b(x) = 1 + ks(1+y) = 1 + ks\left(\frac{1}{|r'|} - x + 1\right). \tag{4.29}$$

The condition $r > -sd$ implies that $a, b \geq 0$. With these notations, we have

$$W_1'(x) = \left(\frac{1}{a} + \frac{1}{b}\right)\left(\frac{1}{|r'|} - 2x - ksxy\left(\frac{1}{a} - \frac{1}{b}\right)\right)$$
$$= \left(\frac{1}{a} + \frac{1}{b}\right)\left(\frac{1}{|r'|} - 2x\right)\left(1 - (ks)^2\frac{xy}{ab}\right).$$

The last identity follows from

$$\frac{1}{a} - \frac{1}{b} = \frac{ks}{ab}\left(\frac{1}{|r'|} - 2x\right).$$

Since $a, b \geq 0$ and $x \in [0, \frac{1}{2|r'|}]$, it suffices to show that

$$ab - (ks)^2xy \geq 0.$$

Using (4.28) and (4.29), we have

$$ab - (ks)^2xy = (1 + ks)\left(1 + \frac{ks}{r}\right).$$

Then the desired statement follows from that $s > -1/d$ and $r > -sd$. We conclude that W_1 is increasing on $[0, \frac{1}{2|r'|}]$.

It remains to show that $W_2(x)$ is increasing on $[0, \frac{1}{2|r'|}]$. Recall the definition of $W_2(x)$ in (4.27), one can check that

$$W_2'(x) = \left(\frac{1}{|r'|} - 2x\right)\left(\frac{1}{a^2} + \frac{1}{b^2}\right) - 2ksxy\left(\frac{1}{a^3} - \frac{1}{b^3}\right)$$
$$= \frac{b-a}{ks}\left(\frac{1}{a^2} + \frac{1}{b^2}\right) - 2ksxy\left(\frac{1}{a^3} - \frac{1}{b^3}\right)$$
$$= \frac{b-a}{ksa^3b^3}T(x),$$

where a and b are defined in (4.28) and (4.29), and

$$T(x) = ab(a^2 + b^2) - 2k^2s^2xy(a^2 + ab + b^2).$$

Since

$$\frac{b-a}{ks} = \frac{1}{|r'|} - 2x \geq 0, \quad x \in [0, \frac{1}{2|r'|}],$$

it suffices to show that $T(x) \geq 0$ for $[0, \frac{1}{2|r'|}]$. Using the identity

$$a'(x)b(x) + a(x)b'(x) = ks(b - a) = -a(x)a'(x) - b(x)b'(x),$$

one can check that

$$T'(x) = ks(a - b)U(x),$$

where

$$U(x) = a^2 + b^2 + 4ab - 2k^2s^2xy.$$

Notice that $U'(x) \equiv 0$, which implies that $U(x)$ is a constant. Since $a, b \geq 0$, we have

$$U(0) = a^2 + b^2 + 4ab > 0.$$

Hence, $T'(x) \leq 0$, i.e., $T(x)$ is decreasing. Therefore, since $a = b$ when $x = \frac{1}{2|r'|}$, we have

$$T(x) \geq T\left(\frac{1}{2|r'|}\right) = 2a^2(a^2 - 3k^2s^2x^2) \quad \text{at } x = \frac{1}{2|r'|}.$$

It suffices to have

$$a^2 \geq 3k^2s^2x^2, \quad x = \frac{1}{2|r'|},$$

which is equivalent to

$$\frac{1}{|r'|} \leq \frac{2}{1+\sqrt{3}}\left(\frac{1}{k|s|} - 1\right).$$

This finishes the proof that every $-g_k(\lambda)/H(\lambda)$ is also increasing on $[0, 1/2]$. Then the numerical value of α in Theorem 4.3 follows from setting $\lambda = 1/2$ in (4.25). □

Remark 4.22 Our optimization argument heavily relies on the fact that $-A(\lambda)/H(\lambda)$ and $-g_k(\lambda)/H(\lambda)$ are monotonically increasing for $\lambda \in [0, 1/2]$. As observed in [27], the monotonicity of $-A(\lambda)/H(\lambda)$ does not depend on the value of r. Numerical examples show that $-g_k(\lambda)/H(\lambda)$, even the whole quantity in (4.26), is not monotone when r is small. This is one of the reasons for the restriction $r > r_0$.

Remark 4.23 Note that the condition $r > -sd$ of Theorem 4.18 can be rewritten as

$$\frac{1}{|r'|} < \left(\frac{1}{d|s|} - 1\right).$$

We do not know whether Theorem 4.3 holds when

$$\frac{2}{1+\sqrt{3}}\left(\frac{1}{d|s|} - 1\right) < \frac{1}{|r'|} < \left(\frac{1}{d|s|} - 1\right).$$

4.4 An Entropic Characterization of s-Concave Densities

Let X and Y be real-valued random variables (possibly dependent) with the identical density f. Cover and Zhang [20] proved that

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

holds for every coupling of X and Y if and only if f is log-concave. This yields an entropic characterization of one-dimensional log-concave densities. We will extend Cover and Zhang's result to Rényi entropies of random vectors with s-concave densities (defined in (4.5)), which particularly include log-concave densities as a special case. This was previously proved in [28] when f is continuous.

Firstly, we introduce some classical variations of convexity and concavity which will be needed in our proof.

Definition 4.24 Let $\lambda \in (0, 1)$ be fixed. A function $f : \mathbb{R}^d \to \mathbb{R}$ with convex support is called almost λ-convex if the following inequality

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \tag{4.30}$$

holds for almost every pair x, y in the domain of f. We say that f is λ-convex if the above inequality holds for every pair x, y in the domain of f. Particularly, for $\lambda = 1/2$, it is usually called mid-convex or Jensen convex. We say that f is convex if f is λ-convex for any $\lambda \in (0, 1)$.

One can define almost λ-concavity, λ-concavity and concavity by reversing inequality (4.30). Adamek [1, Theorem 1] showed that an almost λ-convex function is identical to a λ-convex function except on a set of Lebesgue measure 0. (To apply the theorem there, one can take the ideals $\mathcal{I}_1$ and $\mathcal{I}_2$ as the family of sets with Lebesgue measure 0 in $\mathbb{R}^d$ and $\mathbb{R}^{2d}$, respectively). In general, λ-convexity is not equivalent to convexity, as it is not a strong enough notion to imply continuity, at least not in a logical framework that accepts the axiom of choice. Indeed, counterexamples can be constructed using a Hamel basis for $\mathbb{R}$ as a vector space over $\mathbb{Q}$. However, in the case that f is Lebesgue measurable, a classical result of

Blumberg [6] and Sierpinski [40] (see also [18] in more general setting) shows that λ-convexity implies continuity, and thus convexity.

Theorem 4.25 *Let $s > -1/d$ and we define $r = 1 + s$. Let f be a probability density on $\mathbb{R}^d$. The following statements are equivalent.*

1. *The density f is s-concave.*
2. *For any $\lambda \in (0, 1)$, we have $h_r(\lambda X + (1-\lambda)Y) \le h_r(X)$ for any random vectors X and Y with the identical density f.*
3. *We have $h_r\left(\frac{X+Y}{2}\right) \le h_r(X)$ for any random vectors X and Y with the identical density f.*

Proof We only prove the statement for $s > 0$, or equivalently $r > 1$. The proof for $-1/d < s < 0$, or equivalently $1 - 1/d < r < 1$, is similar and sketched below.

$1 \Longrightarrow 2$: The proof is taken from [28]. We include it for completeness. Let g be the density of $\lambda X + (1-\lambda)Y$. Then we have

$$\begin{aligned} h_r(X) &= \frac{1}{1-r} \log \mathbb{E} f^{r-1}(X) \\ &= \frac{1}{1-r} \log(\lambda \mathbb{E} f^{r-1}(X) + (1-\lambda)\mathbb{E} f^{r-1}(Y)) && (4.31) \\ &\ge \frac{1}{1-r} \log \mathbb{E} f^{r-1}(\lambda X + (1-\lambda)Y) && (4.32) \\ &= \frac{1}{1-r} \log \int_{\mathbb{R}^d} f(x)^{r-1} g(x) dx \\ &\ge \frac{1}{1-r} \log \left(\int_{\mathbb{R}^d} f(x)^r dx \right)^{1-\frac{1}{r}} \left(\int_{\mathbb{R}^d} g(x)^r dx \right)^{\frac{1}{r}} && (4.33) \\ &= \frac{r-1}{r} h_r(X) + \frac{1}{r} h_r(\lambda X + (1-\lambda)Y). \end{aligned}$$

This is equivalent to the desired statement. Identity (4.31) follows from the assumption that X and Y have the same distribution. In inequality (4.32), we use the concavity of f^{r-1} and the fact that $\frac{1}{1-r} \log x$ is decreasing when $r > 1$. Inequality (4.33) follows from Hölder's inequality and the fact that $\frac{1}{1-r} \log x$ is decreasing when $r > 1$. For $1 - 1/d < r < 1$, the statement follows from the same argument in conjunction with the convexity of f^{r-1}, the converse of Hölder's inequality and the fact that $\frac{1}{1-r} \log x$ is increasing when $0 < r < 1$.

$2 \Longrightarrow 3$: Obvious by taking $\lambda = \frac{1}{2}$.

$3 \Longrightarrow 1$: We will prove the statement by contradiction. We first show an example borrowed from Cover and Zhang [20] to illustrate the "mass transferring" argument used in our proof. Consider the density $f(x) = 3/2$ in the intervals $(0, 1/3)$ and $(2/3, 1)$. It is clear that f is not $(r-1)$-concave. The joint distribution of (X, Y) with $Y \equiv X$ is supported on the diagonal line $y = x$. The Radon-Nikodym derivative g with respect to the one-dimensional Lebesgue measure on the line $y = x$ exists and is shown in Fig. 4.1. We remove some "mass" from the diagonal line $y = x$ to

Fig. 4.1 g

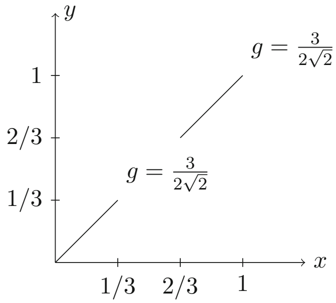
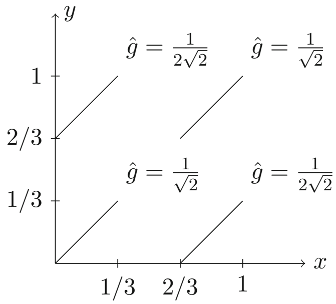


Fig. 4.2 $\hat{g}$



the lines $y = x - 2/3$ and $y = x + 2/3$. The new Radon–Nikodym derivative $\hat{g}$ is shown in Fig. 4.2. Let $(\hat{X}, \hat{Y})$ be a pair of random variables whose joint distribution possesses this new Radon–Nikodym derivative. It is easy to see that $\hat{X}$ and $\hat{Y}$ still have the same density f. But $\hat{X} + \hat{Y}$ is uniformly distributed on $(0, 2)$, and thus $h_r(\hat{X} + \hat{Y}) = \log 2$. One can check that $h_r(2X) = \log(4/3)$.

Now we turn to the general case. Suppose that f is not $(r-1)$-concave, i.e., f^{r-1} is not concave (for $r > 1$). We claim that there exists a set $A \subseteq \mathbb{R}^{2d}$ of positive Lebesgue measure on $\mathbb{R}^{2d}$ such that the inequality

$$2f^{r-1}\left(\frac{x+y}{2}\right) < f^{r-1}(x) + f^{r-1}(y) \tag{4.34}$$

holds for all $(x, y) \in A$. Otherwise, the converse of (4.34) holds for almost every pair (x, y), and thus f^{r-1} is an almost mid-concave function (i.e., 1/2-concave). By Theorem 1 in [1], f^{r-1} is identical to a mid-concave function except on a set of Lebesgue measure 0. Without changing the distribution, we can modify f such that f^{r-1} is mid-concave. Using the equivalence of mid-concavity and concavity (under the Lebesgue measurability), after modification, f^{r-1} is concave, i.e., f is $(r-1)$-concave. This contradicts our assumption. Hence, there exists such a set A with positive Lebesgue measure on $\mathbb{R}^{2d}$. Then there exists y such that (4.34) holds for a set of x with positive Lebesgue measure on $\mathbb{R}^d$. We rephrase this statement in a form suitable for our purpose. There is $x_0 \neq 0$ such that the set

$$\Lambda = \left\{x \in \mathbb{R}^d : 2f(x)^{r-1} < f(x+x_0)^{r-1} + f(x-x_0)^{r-1}\right\} \tag{4.35}$$

has positive Lebesgue measure on $\mathbb{R}^d$. For $\epsilon > 0$, we denote by $\Lambda(\epsilon)$ a ball of radius ϵ whose intersection with Λ has positive Lebesgue measure on $\mathbb{R}^d$. Consider (X, Y) such that $X \equiv Y$, where X and Y have the identical density f. Let $g(x, y)$ be the Radon-Nikodym derivative of (X, Y) with respect to the d-dimensional Lebesgue measure on the "diagonal line" $y = x$. Now we build a new density $\hat{g}$ by translating a small amount of "mass" from "diagonal points" $(x-x_0, x-x_0)$ and $(x+x_0, x+x_0)$ to "off-diagonal points" $(x - x_0, x + x_0)$ and $(x + x_0, x - x_0)$. To be more precise, we define the new joint density $\hat{g}$ as

$$\hat{g}(x, y) = g(x, y)\mathbf{1}_{\{x=y\}} - \sqrt{d/2}\delta(\mathbf{1}_{\{(x-x_0,x-x_0):x\in\Lambda(\epsilon)\}} + \mathbf{1}_{\{(x+x_0,x+x_0):x\in\Lambda(\epsilon)\}}) \\ +\sqrt{d/2}\delta(\mathbf{1}_{\{(x-x_0,x+x_0):x\in\Lambda(\epsilon)\}} + \mathbf{1}_{\{(x+x_0,x-x_0):x\in\Lambda(\epsilon)\}}),$$

where $\delta > 0$ and $\mathbf{1}_S$ is the indicator function of the set S. The function $\hat{g}$ is supported on the "diagonal line" $y = x$ and "off-diagonal segments" $\{(x - x_0, x + x_0) : x \in \Lambda(\epsilon)\}$ and $\{(x + x_0, x - x_0) : x \in \Lambda(\epsilon)\}$, which are disjoint for sufficiently small $\epsilon > 0$. (This is similar to Fig. 4.2.) When $\delta > 0$ is small enough, $\hat{g}(x, y)$ is non-negative everywhere. Furthermore, our construction preserves the "total mass". Hence, the function $\hat{g}(x, y)$ is indeed a probability density with respect to the d-dimensional Lebesgue measure on the "diagonal line" and two "off-diagonal segments". Let $(\hat{X}, \hat{Y})$ be a pair with the joint density $\hat{g}(x, y)$. The marginals $\hat{X}$ and $\hat{Y}$ have the same distribution as that of X, since the "positive mass" on "off-diagonal points" complements the "mass deficit" on "diagonal points" when we project in the x and y directions. We claim that $\frac{\hat{X}+\hat{Y}}{2}$ has larger entropy than $\hat{X}$. One can check that the density of $\frac{\hat{X}+\hat{Y}}{2}$ is

$$\hat{f}(x) = f(x) + \delta(2\mathbf{1}_{\Lambda(\epsilon)} - \mathbf{1}_{\Lambda(\epsilon)+x_0} - \mathbf{1}_{\Lambda(\epsilon)-x_0}).$$

Let Ω denote the union of $\Lambda(\epsilon)$, $\Lambda(\epsilon) + x_0$ and $\Lambda(\epsilon) - x_0$. Then we have

$$h_r\left(\frac{\hat{X}+\hat{Y}}{2}\right) = \frac{1}{1-r}\log\left(\int_\Omega \hat{f}(x)^r dx + \int_{\Omega^c} f^r(x)dx\right). \tag{4.36}$$

Since $x_0 \neq 0$, for $\epsilon > 0$ small enough, Ω is the union of disjoint translates of $\Lambda(\epsilon)$. When $\delta > 0$ is sufficiently small, we have

$$\int_\Omega \hat{f}(x)^r dx = \int_{\Lambda(\epsilon)} \left[(f(x) + 2\delta)^r + (f(x + x_0) - \delta)^r + (f(x - x_0) - \delta)^r\right] dx$$

$$< \int_{\Lambda(\epsilon)} \left[f(x)^r + f(x + x_0)^r + f(x - x_0)^r\right] dx \tag{4.37}$$

$$= \int_\Omega f(x)^r dx, \tag{4.38}$$

where inequality (4.37) follows from the observation that for $x \in \Lambda(\epsilon) \subset \Lambda$ (see (4.35)) the derivative of the integrand at $\delta = 0$ is

$$r[2f(x)^{r-1} - f(x-x_0)^{r-1} - f(x+x_0)^{r-1}] < 0. \quad (4.39)$$

Since $r > 1$, (4.36) together with (4.38) implies that

$$h_r\left(\frac{\hat{X}+\hat{Y}}{2}\right) > \frac{1}{1-r}\log\left(\int_\Omega f(x)^r dx + \int_{\Omega^c} f(x)^r dx\right) = h(X) = h(\hat{X}).$$

This is contradictory to our assumption. Hence, f has to be $(r-1)$-concave. For $1 - 1/d < r < 1$, we redefine the set Λ by reversing inequality (4.35), and inequality (4.37) will be also reversed. We will arrive at the same conclusion. □

Remark 4.26 The proof of $1 \Longrightarrow 2$ is an immediate consequence of Theorem 3.36 in [30]. The theorem there draws heavily on the ideas of [42], where a related study, deriving the Schur convexity of Rényi entropies under the assumption of exchangeability and s-concavity of the random variables, generalizing Yu's results in [43] on the entropies of sums of i.i.d. log-concave random variables. Although we state Theorem 4.25 for two random vectors, the argument also works for more than two random vectors. Hence, it implies the seemingly stronger Theorem 4.4.

As an immediate consequence of Theorem 4.25, we have the following reverse Rényi EPI for random vectors with the same distribution.

Corollary 4.27 *Let $s > -1/d$ and let $r = 1 + s$. Let X and Y be (possibly dependent) random vectors in $\mathbb{R}^d$ with the same density f being s-concave. Then we have*

$$N_r(X+Y) \le 4N_r(X).$$

References

1. M. Adamek, Almost λ-convex and almost Wright-convex functions. Math. Slovaca **53**(1), 67–73 (2003)
2. K. Ball, P. Nayar, T. Tkocz, A reverse entropy power inequality for log-concave random vectors. Stud. Math. **235**(1), 17–30 (2016)
3. A.R. Barron, Entropy and the central limit theorem. Ann. Probab. **14**, 336–342 (1986)
4. W. Beckner, Inequalities in Fourier analysis. Ann. Math. **102**(1), 159–182 (1975)
5. R.N. Bhattacharya, R. Ranga Rao, *Normal Approximation and Asymptotic Expansions*. (Wiley/Society for Industrial and Applied Mathematics, Hoboken/Philadelphia, 2010/1976)
6. H. Blumberg, On convex functions. Trans. Am. Math. Soc. **20**, 40–44 (1919)
7. S.G. Bobkov, G.P. Chistyakov, Bounds for the maximum of the density of the sum of independent random variables. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. **408**(18), 62–73, 324 (2012)

8. S.G. Bobkov, G.P. Chistyakov, Entropy power inequality for the Rényi entropy. IEEE Trans. Inform. Theory **61**(2), 708–714 (2015)
9. S. Bobkov, M. Madiman, Reverse Brunn-Minkowski and reverse entropy power inequalities for convex measures. J. Funct. Anal. **262**, 3309–3339 (2012)
10. S.G. Bobkov, A. Marsiglietti, Variants of the entropy power inequality. IEEE Trans. Inform. Theory **63**(12), 7747–7752 (2017)
11. S.G. Bobkov, A. Marsiglietti, Asymptotic behavior of Rényi entropy in the central limit theorem (2018). Preprint. arXiv:1802.10212
12. S.G. Bobkov, G.P. Chistyakov, F. Götze, Rényi divergence and the central limit theorem. Ann. Probab. **47**(1), 270–323 (2019)
13. C. Borell, Convex measures on locally convex spaces. Ark. Mat. **12**, 239–252 (1974)
14. C. Borell, Convex set functions in d-space. Period. Math. Hungar. **6**(2), 111–136 (1975)
15. H.J. Brascamp, E.H. Lieb, Best constants in Young's inequality, its converse, and its generalization to more than three functions. Adv. Math. **20**(2), 151–173 (1976)
16. S. Brazitikos, A. Giannopoulos, P. Valettas, B.-H. Vritsiou, *Geometry of isotropic convex bodies*. AMS-Mathematical Surveys and Monographs, vol. 196 (American Mathematical Society, Providence, 2014)
17. H. Busemann, A theorem on convex bodies of the Brunn-Minkowski type. Proc. Natl. Acad. Sci. U. S. A. **35**, 27–31 (1949)
18. A. Chademan, F. Mirzapour, Midconvex functions in locally compact groups. Proc. Am. Math. Soc. **127**, 2961–2968 (1999)
19. M.H.M. Costa, A new entropy power inequality. IEEE Trans. Inform. Theory **31**(6), 751–760 (1985)
20. T.M. Cover, Z. Zhang, On the maximum entropy of the sum of two dependent random variables. IEEE Trans. Inform. Theory **40**(4), 1244–1246 (1994)
21. A. Dembo, T.M. Cover, J.A. Thomas, Information-theoretic inequalities. IEEE Trans. Inform. Theory **37**(6), 1501–1518 (1991)
22. M. Fradelizi, J. Li, M. Madiman, Concentration of information content for convex measures (2015). Preprint. arXiv:1512.01490
23. R.J. Gardner, A. Giannopoulos, p-cross-section bodies. Indiana Univ. Math. J. **48**(2), 593–614 (1999)
24. B.V. Gnedenko, A.N. Kolmogorov, *Limit distributions for sums of independent random variables*. Translated from the Russian, annotated, and revised by K. L. Chung. With appendices by J. L. Doob and P. L. Hsu. Revised edition. (Addison-Wesley Publishing, Reading, 1968)
25. W. Hoeffding, Probability inequalities for sums of bounded random variables. J. Am. Stat. Assoc. **58**, 13–30 (1963)
26. O. Johnson, *Information theory and the central limit theorem* (Imperial College Press, London, 2004)
27. J. Li, Rényi entropy power inequality and a reverse. Stud. Math. **242**, 303–319 (2018)
28. J. Li, J. Melbourne, Further investigations of the maximum entropy of the sum of two dependent random variables, in *Proceedings of IEEE International Symposium on Information Theory*, Vail (2018), pp. 1969–1972
29. E.H. Lieb, Proof of an entropy conjecture of Wehrl. Commun. Math. Phys. **62**(1), 35–41 (1978)
30. M. Madiman, J. Melbourne, P. Xu, Forward and reverse entropy power inequalities in convex geometry. Convexity Concentration **161**, 427–485 (2017)
31. M. Madiman, J. Melbourne, P. Xu, Rogozin's convolution inequality for locally compact groups (2017). Preprint. arXiv:1705.00642
32. M. Madiman, J. Melbourne, P. Xu, Infinity-Rényi entropy power inequalities, in *2017 IEEE International Symposium on Information Theory (ISIT)*, Aachen (2017), pp. 2985–2989
33. A. Marsiglietti, J. Melbourne, On the entropy power inequality for the Rényi entropy of order [0, 1]. IEEE Trans. Inform. Theory **65**, 1387–1396 (2018). https://doi.org/10.1109/TIT.2018.2877741

34. A. Marsiglietti, J. Melbourne, A Rényi entropy power inequality for log-concave vectors and parameters in [0, 1], in *Proceedings of IEEE International Symposium on Information Theory*, Vail (2018), pp. 1964–1968
35. V.D. Milman, Inégalité de Brunn-Minkowski inverse et applications à la théorie locale des espaces normés. C. R. Acad. Sci. Paris Sér. I Math. **302**(1), 25–28 (1986)
36. E. Ram, I. Sason, On Rényi entropy power inequalities. IEEE Trans. Inform. Theory **62**(12), 6800–6815 (2016)
37. O. Rioul, Rényi entropy power inequalities via normal transport and rotation. Entropy **20**(9), 641 (2018)
38. G. Savaré, G. Toscani, The concavity of Rényi entropy power. IEEE Trans. Inform. Theory **60**(5), 2687–2693 (2014)
39. C.E. Shannon, A mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–423 623–656 (1948)
40. W. Sierpinski, Sur les fonctions convexes mesurables. Fund. Math. **1**, 125–129 (1920)
41. A.J. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon. Inf. Control. **2**, 101–112 (1959)
42. P. Xu, J. Melbourne, M. Madiman, Reverse entropy power inequalities for s-concave densities, in *Proceedings of International Symposium on Information Theory*, Barcelona (2016), pp. 2284–2288
43. Y. Yu, Letter to the editor: on an inequality of Karlin and Rinott concerning weighted sums of i.i.d. random variables. Adv. Appl. Probab. **40**(4), 1223–1226 (2008)