Entropic Central Limit Theorem for Rényi Entropy

Jiange Li
Einstein Institute of Mathematics
The Hebrew University of Jerusalem
Jerusalem, 9190401, Israel
Email: jiange.li@mail.huji.ac.il

Arnaud Marsiglietti
Department of Mathematics
University of Florida
Gainesville, FL 32611, USA
Email: a.marsiglietti@ufl.edu

James Melbourne
Electrical and Computer Engineering
University of Minnesota
Minneapolis, MN 55455, USA
Email: melbo013@umn.edu

***Abstract*—We establish a central limit theorem for Rényi entropies when the Rényi parameters belong to $(0,1)$ for a large class of random vectors. This complements a celebrated result of Barron (1986). As an application, we show that a general Rényi entropy power inequality fails when the Rényi parameter is in $(0,1)$.**

***Index Terms*—Rényi entropy; entropy power inequality; entropic central limit theorem**

I. Introduction

The Central Limit Theorem (henceforth, CLT) is a fundamental result in probability theory and statistics. It also has a plethora of applications in applied sciences. The CLT has close connections with information theory via the Entropy Power Inequality (henceforth, EPI) of Shannon [25] and Stam [26].

Let X be a random vector in $\mathbb{R}^d$ with density f. The Shannon differential entropy $h(X)$ is defined as

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

The entropy power $N(X)$ is defined as

$$N(X) = e^{2h(X)/d}. \tag{2}$$

Shannon-Stam's EPI states that

$$N(X+Y) \geq N(X) + N(Y) \tag{3}$$

holds for arbitrary independent random vectors X and Y in $\mathbb{R}^d$. Let $\{X_n\}_{n\geq 1}$ be a sequence of independent and 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

$$Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}. \tag{4}$$

By induction and the homogeneity of $N(\cdot)$, the EPI (3) implies that the dyadic sequence $\{h(Z_{2^n})\}_{n\geq 1}$ is non-decreasing. It is well known that Gaussian maximizes the Shannon entropy when the covariance matrix is fixed (see, e.g., [10]). Hence, we have $h(Z_{2^n}) \leq h(Z)$, where Z is a centered Gaussian vector with the same covariance matrix as X_1. This implies the convergence of the dyadic sequence $\{h(Z_{2^n})\}_{n\geq 1}$. A celebrated result of Barron [2] states that

$$h(Z_n) \to h(Z), \tag{5}$$

as $n \to +\infty$, which strengthens the classical CLT that the sequence $\{Z_n\}_{n\geq 1}$ converges to Z in distribution. The entropic CLT (5) along with the monotonicity of $\{h(Z_{2^n})\}_{n\geq 1}$ and maximization of entropy by Gaussians can be seen as an analogue of the second law of thermodynamics (see, e.g., [13]). Furthermore, it was proven in [1] that the sequence $\{h(Z_n)\}_{n\geq 1}$ is actually monotone at each step, and not only along dyadic steps, see also [17].

There has been recent success in extending the Shannon EPI (3) to the Rényi setting (see [4], [5], [7], [14], [18], [19], [21], [23], [24]) but few results are known about CLT for r-Rényi entropy (see [8] for $r > 1$ in dimension 1; see also [6] for convergence in Rényi divergence, which is not equivalent to convergence in Rényi entropy unless $r = 1$). The aim of this paper is to extend the above Shannon entropic CLT to the Rényi entropic setting, with particular interest in application to Rényi EPI. For $r \geq 0$, the r-Rényi entropy of a random vector X in $\mathbb{R}^d$ with density f is defined as

$$h_r(X) = \frac{1}{1-r}\log\int_{\mathbb{R}^d} f(x)^r dx. \tag{6}$$

For $r \in \{0, 1, \infty\}$, the definition is understood in the limiting sense, and $h_1(X)$ is the Shannon differential entropy. The r-Rényi entropy power $N_r(X)$ is defined as

$$N_r(X) = e^{2h_r(X)/d}. \tag{7}$$

This paper is organized as follows. The next section is dedicated to the convergence of $h_r(Z_n)$. For $r > 1$ convergence is fully characterized for random vectors in $\mathbb{R}^d$, while for $r \in (0,1)$ sufficient conditions with application to Rényi EPI are explored. Explicitly, convergence is proven for log-concave random vectors and random vectors with radially symmetric unimodal densities and compact support. The last section presents applications of the Rényi entropic CLT. Most significantly, we show that a general r-Rényi EPI fails when $r \in (0,1)$. The reader is directed to the full paper expanding on this work [15].

II. Central Limit Theorem for Rényi entropy

A fundamental tool in establishing various central limit theorems is the characteristic function. Recall that the characteristic function of a random vector X in $\mathbb{R}^d$ is defined by

$$\varphi_X(t) = \mathbb{E}[e^{i\langle t,X\rangle}], \quad t \in \mathbb{R}^d. \tag{8}$$

Before providing sufficient conditions for convergence in r-Rényi entropy, with $r \in (0,1)$, we first extend [8, Theorem 1.1] to the higher dimensional setting.

Theorem 2.1: Let $r > 1$. Let $X_1, \dots, X_n$ be i.i.d. centered random vectors in $\mathbb{R}^d$. Let us define Z_n as in (4) and denote by ρ_n the density of Z_n. Then, the following statements are equivalent.

1) $h_r(Z_n) \to h_r(Z)$, where Z is a Gaussian random vector with mean 0 and 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: $1 \implies 2$: Assume that $h_r(Z_n) \to h_r(Z)$ as $n \to +\infty$. Then, there exists an integer n_0 such that

$$h_r(Z) - 1 < h_r(Z_{n_0}) < h_r(Z) + 1. \tag{9}$$

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

$2 \implies 3$: Assume that $h_r(Z_{n_0})$ is finite for some integer n_0. Then, Z_{n_0} has a density ρ_{n_0} which is in $L^r(\mathbb{R}^d)$, and thus Z_n has a density $\rho_n \in L^r$ for any $n \geq n_0$ by the convolution structure of Z_n.

Case 1: If $r \geq 2$, then $\rho_n \in L^2(\mathbb{R}^d)$. Hence by Plancherel identity, $\varphi_{Z_n} \in L^2(\mathbb{R}^d)$. It follows that

$$\int_{\mathbb{R}^d} |\varphi_{Z_n}(t)|^2 \, dt = \int_{\mathbb{R}^d} |\varphi_{X_1}\left(t/\sqrt{n}\right)|^{2n} \, dt < +\infty. \tag{10}$$

We deduce that for $\nu = 2n_0$,

$$\int_{\mathbb{R}^d} |\varphi_{X_1}(t)|^\nu \, dt < +\infty. \tag{11}$$

Case 2: If $r \in (1, 2)$, then by the Hausdorff-Young inequality,

$$\|\varphi_{Z_n}\|_{L^{r'}} \leq \frac{1}{(2\pi)^{d/r'}} \|\rho_n\|_{L^r}, \tag{12}$$

where r' is the conjugate of r. Hence, for $\nu = r' n_0$,

$$\int_{\mathbb{R}^d} |\varphi_{X_1}(t)|^\nu \, dt < +\infty. \tag{13}$$

$3 \implies 4$: 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 [11]), which is valid in arbitrary dimension (see [3]). In particular,

$$\lim_{n \to +\infty} \sup_{x \in \mathbb{R}^d} |\rho_n(x) - \phi_\Sigma(x)| = 0, \tag{14}$$

where ϕ_Σ denotes the density of a Gaussian random vector with mean 0 and 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 $n \geq n_0$.

$4 \implies 1$: Since ρ_{n_0} is bounded, then $\rho_{n_0} \in L^2$, and we deduce by Plancherel identity that $\int_{\mathbb{R}^d} |\varphi_{X_1}(t)|^\nu \, dt < +\infty$ for $\nu = 2n_0$. Hence (14) holds and there exists $M > 0$ such that $\rho_n \leq M$ for all $n \geq n_0$. Let us show that $\int \rho_n^r \to \int \phi_\Sigma^r$ as $n \to +\infty$, where ϕ_Σ denotes the density of a Gaussian random vector with mean 0 and same covariance matrix as X_1. From the central limit theorem, there exists $T > 0$ such that for all n large enough,

$$\int_{|x|>T} \rho_n(x) dx < \varepsilon, \qquad \int_{|x|>T} \phi_\Sigma(x) dx < \varepsilon. \tag{15}$$

Hence,

$$\int_{|x|>T} \rho_n^r(x) dx \leq M^{r-1} \int_{|x|>T} \rho_n(x) dx < M^{r-1}\varepsilon, \tag{16}$$

and similarly for $\int_{|x|>T} \phi_\Sigma^r$. Hence, for any $\delta > 0$, there exists $T > 0$ such that for all n large enough,

$$\left| \int_{|x|>T} \rho_n^r(x) dx - \int_{|x|>T} \phi_\Sigma^r(x) dx \right| < \delta. \tag{17}$$

On the other hand, by (14), for all $T > 0$, the function $\rho_n^r(x) 1_{\{|x| \leq T\}}$ converges everywhere to $\phi_\Sigma^r(x) 1_{\{|x| \leq T\}}$ as $n \to +\infty$. Since $\rho_n^r(x) 1_{\{|x| \leq T\}}$ is dominated by the integrable function $M^r 1_{\{|x| \leq T\}}$, one may use the Lebesgue dominated theorem to conclude that

$$\lim_{n \to +\infty} \left| \int_{|x| \leq T} \rho_n^r(x) dx - \int_{|x| \leq T} \phi_\Sigma^r(x) dx \right| = 0. \tag{18}$$

■

Remark 2.1: When $r \in (0, 1)$, the statement of Theorem 2.1 fails since it is possible to find a bounded density ρ such that $\int \rho(x)^r \, dx = +\infty$ (e.g., Cauchy-type distributions). In particular, for random variable X_1 with such density, the implication $4 \implies 2$ (and thus $4 \implies 1$) in Theorem 2.1 does not necessarily hold when $r \in (0, 1)$ since by Jensen inequality $h_r(Z_n) \geq h_r(X_1/\sqrt{n}) = \infty$, for all $n \geq 1$. It was noted by Barron [2] that in the Shannon entropy case $r = 1$, the implication $1 \implies 4$ does not necessarily hold.

In the next results, we provide sufficient conditions for a CLT to hold for r-Rényi entropy, with $r \in (0, 1)$, for different classes of random vectors in $\mathbb{R}^d$. Recall the definition of Z_n in (4).

Theorem 2.2: Let $r \in (0, 1)$. Let $\{X_n\}_{n \geq 1}$ be a sequence of i.i.d. centered log-concave random vectors in $\mathbb{R}^d$. Then, $h_r(Z_n) < +\infty$ for all $n \geq 1$, and

$$h_r(Z_n) \to h_r(Z),$$

where Z is a Gaussian random vector with mean 0 and same covariance matrix as X_1.

Proof: Since log-concavity is preserved under independent sum [22], 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

$$\rho_n(x) \leq e^{-a_n|x| + b_n}, \tag{19}$$

for all $x \in \mathbb{R}^d$, for some constants $a_n > 0$, $b_n \in \mathbb{R}$ possibly depending on the dimension (see, e.g., [9]). Hence, for all $n \geq 1$,

$$\int_{\mathbb{R}^d} \rho_n(x)^r \, dx \leq \int_{\mathbb{R}^d} e^{-r(a_n|x| + b_n)} \, dx < +\infty. \tag{20}$$

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

The boundedness of ρ_n implies that (14) holds, and thus there exists an integer n_0 such that for all $n \geq n_0$,

$$\rho_n(0) > \frac{1}{2} \phi_\Sigma(0), \tag{21}$$

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 $x \in \mathbb{R}^d$,

$$\begin{aligned}\rho_n(rx) &= \rho_n((1-r)0 + rx) \\ &\geq \rho_n(0)^{1-r}\rho_n(x)^r \\ &\geq \frac{1}{2^{1-r}}\phi_\Sigma(0)^{1-r}\rho_n(x)^r.\end{aligned} \tag{22}$$

Hence, for all $T > 0$,

$$\int_{|x|>T} \rho_n(x)^r\, dx \leq \frac{2^{1-r}}{\phi_\Sigma(0)^{1-r}} \int_{|x|>T} \rho_n(rx)\, dx \tag{23}$$

$$= \frac{2^{1-r}}{r^d \phi_\Sigma(0)^{1-r}} \mathbb{P}(|Z_n| > rT) \tag{24}$$

$$\leq \frac{1}{T^2} \frac{2^{1-r}\mathbb{E}[|X_1|^2]}{r^{d+2}\phi_\Sigma(0)^{1-r}}, \tag{25}$$

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

$$\mathbb{E}[|Z_n|^2] = \frac{\mathbb{E}[|X_1|^2] + \cdots + \mathbb{E}[|X_n|^2]}{n} = \mathbb{E}[|X_1|^2]. \tag{26}$$

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

$$\int_{|x|>T} \rho_n^r(x)dx < \varepsilon, \qquad \int_{|x|>T} \phi_\Sigma^r(x)dx < \varepsilon, \tag{27}$$

and hence

$$\left| \int_{|x|>T} \rho_n^r(x)dx - \int_{|x|>T} \phi_\Sigma^r(x)dx \right| < \varepsilon. \tag{28}$$

On the other hand, from (14), we conclude as in the proof of Theorem 2.1 that for all $T > 0$,

$$\lim_{n\to+\infty} \left| \int_{|x|\leq T} \rho_n^r(x)dx - \int_{|x|\leq T} \phi_\Sigma^r(x)dx \right| = 0. \tag{29}$$

■

Next, we provide a convergence result for the more general class of unimodal distributions under additional symmetry assumptions. First, we need the following stability result.

Proposition 2.1: The class of unimodal spherically symmetric random variables is stable under convolution.

Proof: Suppose that f_i are densities such that $f_i(Tx) = f_i(x)$ for an orthogonal map T and $|x| \leq |y|$ implies $f_i(x) \geq f_i(y)$. By the layer cake decomposition, we write

$$f_i(x) = \int_0^\infty \mathbb{1}_{\{(u,v): f_i(u) > v\}}(x, \lambda)d\lambda. \tag{30}$$

After applying Fubini-Tonelli,

$$f_1 \star f_2(x) \tag{31}$$

$$= \int_{\mathbb{R}^d} f_1(x-y)f_2(y)dy$$

$$= \int_0^\infty \int_0^\infty \left(\int_{\mathbb{R}^d} \mathbb{1}_{\{(u,v):f_1(u)>v\}}(x-y, \lambda_1) \right. \tag{32}$$

$$\left. \times\, \mathbb{1}_{\{(u,v):f_2(u)>v\}}(y, \lambda_2)dy \right) d\lambda_1 d\lambda_2. \tag{33}$$

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

$$L_{\lambda_i} = \{u : f_i(u) > \lambda_i\} \tag{34}$$

is an origin symmetric ball. Thus we can write the integrand in (33) as

$$\int_{\mathbb{R}^d} \mathbb{1}_{L_{\lambda_1}}(x-y)\mathbb{1}_{L_{\lambda_2}}(y)dy = \mathbb{1}_{L_{\lambda_1}} \star \mathbb{1}_{L_{\lambda_2}}(x). \tag{35}$$

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

$$\mathbb{1}_{L_{\lambda_1}} \star \mathbb{1}_{L_{\lambda_2}}(x) \geq \mathbb{1}_{L_{\lambda_1}} \star \mathbb{1}_{L_{\lambda_2}}(y). \tag{36}$$

Then we can finish the proof by integrating both sides of the above inequality. ■

Let us establish large deviation and pointwise inequalities for radially symmetric unimodal densities with compact support.

Theorem 2.3 (Hoeffding [12]): For independent random variables X_i with zero mean, bounded in (a_i, b_i), $i = 1, \ldots, n$, one has for all $T > 0$,

$$\mathbb{P}\left(\sum_{i=1}^n X_i > T\right) \leq \exp\left\{-\frac{2T^2}{\sum_{i=1}^n (b_i - a_i)^2}\right\}. \tag{37}$$

Lemma 2.1: For centered independent random vectors X_i in $\mathbb{R}^d$ satisfying $\mathbb{P}(|X_i| > R) = 0$, $i = 1, \ldots, n$, for some $R > 0$, one has for all $T > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right| > T\right) \leq 2d \exp\left\{-\frac{T^2}{2d^2R^2}\right\}. \tag{38}$$

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

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right| > T\right) \tag{39}$$

$$\leq \mathbb{P}\left(\bigcup_{j=1}^d \left\{|X_{1,j} + \cdots + X_{n,j}| > \frac{T\sqrt{n}}{d}\right\}\right) \tag{40}$$

$$\leq \sum_{j=1}^d \mathbb{P}\left(|X_{1,j} + \cdots + X_{n,j}| > \frac{T\sqrt{n}}{d}\right) \tag{41}$$

$$\leq 2d \exp\left(-\frac{T^2}{2d^2R^2}\right), \tag{42}$$

where inequality (40) follows from the pigeon-hole principle, (41) from a union bound, and (42) follows from applying Hoeffding's inequality (Theorem 2.3) to $X_{1,j} + \cdots + X_{n,j}$ and $(-X_{1,j}) + \cdots + (-X_{n,j})$. ■

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

Corollary 2.1: If X_i are i.i.d. with radially symmetric unimodal density function supported on the Euclidean ball $B_R = \{x : |x| \leq R\}$ for some $R > 0$, then letting ρ_n denote density of the normalized sum $(X_1 + \cdots + X_n)/\sqrt{n}$, there exists $c_d > 0$ such that for $|x| > 2$,

$$\rho_n(x) \le c_d \exp\left\{-\frac{(|x|-1)^2}{2d^2R^2}\right\}. \tag{43}$$

Proof: Stating Lemma 2.1 in terms of ρ_n, we have

$$\int_{|y|>T} \rho_n(w)dw \le 2d \exp\left\{-\frac{T^2}{2d^2R^2}\right\}. \tag{44}$$

Since the class of radially symmetric unimodal random variables is stable under independent summation by Proposition 2.1, ρ_n is radially symmetric and unimodal, so that

$$\rho_n(x) \le \frac{\int_{B_{|x|}\setminus B_{|x|-1}} \rho_n(w)dw}{\mathrm{Vol}(B_{|x|}\setminus B_{|x|-1})} \tag{45}$$

$$\le \frac{\int_{|y|\ge|x|-1} \rho_n(w)dw}{(2^d-1)\omega_d} \tag{46}$$

where ω_d is the volume of the unit ball. Note that

$$\mathrm{Vol}(B_{|x|}\setminus B_{|x|-1}) = (|x|^d - (|x|-1)^d)\omega_d \ge (2^d-1)\omega_d, \tag{47}$$

since $t \mapsto t^d - (t-1)^d$ is increasing, so that (46) follows. Now applying (44) we have

$$\rho_n(x) \le \frac{\int_{|y|\ge|x|-1} \rho_n(w)dw}{(2^d-1)\omega_d} \tag{48}$$

$$\le \frac{2d}{(2^d-1)\omega_d} \exp\left\{-\frac{(|x|-1)^2}{2d^2R^2}\right\} \tag{49}$$

and our result holds with

$$c_d = \frac{2d}{(2^d-1)\omega_d}. \tag{50}$$

■

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

Theorem 2.4: Let $r \in (0,1)$. Let $\{X_n\}_{n\ge1}$ be a sequence of i.i.d. random vectors in $\mathbb{R}^d$ with a radially symmetric unimodal density with compact support. Then,

$$\lim_{n\to\infty} N_r\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right) = N_r(Z) \tag{51}$$

where Z is a Gaussian random vector with mean 0 and same covariance matrix as X_1.

Proof: Let us denote by ρ_n the density of the normalized sum $(X_1 + \cdots + X_n)/\sqrt{n}$. Since $\rho_1 \in L^1$, it follows that ρ_n, $n \ge 2$, are continuous, and since ρ_n are, in addition, radially symmetric unimodal densities by Proposition 2.1, then ρ_n, $n \ge 2$, are bounded. Hence, the point 3) of Theorem 2.1 holds with $\nu = 2$, and one may thus apply (14) together with Lebesgue dominated convergence to conclude that for all $T > 0$,

$$\lim_{n\to+\infty} \left| \int_{|x|\le T} \rho_n^r(x)dx - \int_{|x|\le T} \phi_\Sigma^r(x)dx \right| = 0. \tag{52}$$

On the other hand, by Corollary 2.1, for all $\varepsilon > 0$, one may choose $T > 0$ such that for all $n \ge 1$,

$$\int_{|x|>T} \rho_n^r(x)dx < \varepsilon, \qquad \int_{|x|>T} \phi_\Sigma^r(x)dx < \varepsilon, \tag{53}$$

and hence

$$\left| \int_{|x|>T} \rho_n^r(x)dx - \int_{|x|>T} \phi_\Sigma^r(x)dx \right| < \varepsilon. \tag{54}$$

■

III. No Rényi EPI of order $r \in (0,1)$

Definition 3.1: For $r \in [0,\infty]$, define c_r as the largest number such that for any finite sequence of independent random vectors X_i in $\mathbb{R}^d$, the inequality

$$N_r(X_1 + \cdots + X_n) \ge c_r \sum_{i=1}^n N_r(X_i) \tag{55}$$

holds.

Theorem 3.1: For $r \in (0,1)$, the constant c_r defined in (55) satisfies

$$c_r = 0. \tag{56}$$

Theorem 3.1 affirms a striking difference between Rényi EPI of parameter $r \ge 1$ and $r \in (0,1)$. Indeed, it was shown in [5] that for $r \ge 1$, one has

$$c_r \ge \frac{1}{e} r^{\frac{1}{r-1}}. \tag{57}$$

Theorem 3.1 can be reformulated as follows.

Theorem 3.2: For any $r \in (0,1)$ and $\varepsilon > 0$, there exist independent random vectors $X_1, \ldots, X_n$ in $\mathbb{R}^d$, for some $d \ge 1$ and $n \ge 2$, such that

$$N_r(X_1 + \cdots + X_n) < \varepsilon \sum_{i=1}^n N_r(X_i). \tag{58}$$

The motivating observation for this line of argument is the fact that for $r \in (0,1)$, there exist variables with finite variance and infinite r-Rényi entropy. One might anticipate that this could contradict the existence of an r-Rényi EPI, as the central limit theorem forces i.i.d. sums to become "more Gaussian". Heuristically, one anticipates for large n, and X_i drawn from such a distribution, that $N_r(X_1 + \cdots + X_n)/n = N_r((X_1 + \cdots + X_n)/\sqrt{n})$ should approach $N_r(Z)$, where Z is a Gaussian with variance equal to X_i's, while $\sum N_r(X_i)/n = N_r(X_1)$ is infinite.

Proof of Theorem 3.1: Let us consider the following density

$$f(x) = f_{R,p,d}(x) = C_R(1+|x|)^{-p}\mathbb{1}_{B_R}(x), \quad x \in \mathbb{R}^d, \tag{59}$$

with $p, R > 0$ and C_R implicitly determined to make f a density. Note that f is bounded, unimodal, and radially symmetric. Thus its covariance matrix is a multiple of the identity, i.e., $\sigma_R^2 I$ for some $\sigma_R > 0$. Computing in spherical coordinates one can easily see 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 $p > d/r$.

Now fix $r \in (0,1)$ and take the dimension to be $d^* = \min\{d \in \mathbb{N} : d > 2r/(1-r)\}$, and $p \in (d^*+2, d^*/r)$. In this case, the limit 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 radially symmetric unimodal density with compact support, one may apply Theorem 2.4 to conclude that

$$\lim_{n\to\infty} N_r\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right) = \sigma_R^2 N_r(Z_{Id}), \tag{60}$$

where Z_{Id} is the standard d-dimensional Gaussian. Notice that

$$\lim_{R\to\infty} N_r(X_1) = \infty, \tag{61}$$

while

$$\lim_{R\to\infty} \sigma_R = \sigma_\infty < \infty. \tag{62}$$

Given $M > 0$, we can take R large enough such that $N_r(X_1) \geq M$, and $|\sigma_R^2 - \sigma_\infty^2| \leq 1$. Then we can take n large enough such that

$$N_r\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right) \leq (\sigma_\infty^2 + 2) N_r(Z_{Id}). \tag{63}$$

We conclude that for the inequality (55) to hold. We must have

$$c_r \leq \frac{(\sigma_\infty^2+2)N_r(Z_{Id})}{M} \tag{64}$$

for all $M > 0$. Taking $M \to \infty$ this can only hold if $c_r = 0$. ■

Remark 3.1: The counterexample we have built in the proof of Theorem 3.1 is a random vector in $\mathbb{R}^d$ with an s-concave density with $s < 0$ and $|s| > \frac{r}{d}$. In [16], where the terminology of s-concave density is recalled, we prove that for $|s| < \frac{r}{d}$, a Rényi EPI does hold for fixed $r \in (0,1)$, hence complementing the above negative result and extending the following first announced in [20].

Theorem 3.3 ([21]): For $r \in (0,1)$, there exists $C(r) > 0$ such that for all $X_1, \dots, X_n$ independent log-concave random vectors in $\mathbb{R}^d$,

$$N_r(X_1+\cdots+X_n) \geq C(r) \sum_{i=1}^n N_r(X_i). \tag{65}$$

In particular one can take $C(r) = e\, r^{\frac{1}{1-r}}$.

By applying Theorem 2.2 to X_i distributed according to a Laplace distribution, which is log-concave, it follows that the optimal value $C(r)$ satisfying (65) for all log-concave random vectors verifies $C(r) \leq \pi r^{\frac{1}{1-r}}$.

Acknowledgement

James Melbourne's research supported by NSF grants CNS 1544721 and ECCS 1809194.

References

[1] S. Artstein, K. M. Ball, F. Barthe, and A. Naor. Solution of Shannon's problem on the monotonicity of entropy. *J. Amer. Math. Soc.*, 17(4):975–982 (electronic), 2004.

[2] A. R. Barron. Entropy and the central limit theorem. *Ann. Probab.*, 14:336–342, 1986.

[3] R. N. Bhattacharya and R. Ranga Rao. Normal approximation and asymptotic expansions. *John Wiley & Sons, Inc. Also: Soc. for Industrial and Appl. Math., Philadelphia, 2010*, 1976.

[4] S. G. Bobkov and 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. (POMI)*, 408(Veroyatnost i Statistika. 18):62–73, 324, 2012.

[5] S. G. Bobkov and G. P. Chistyakov. Entropy power inequality for the Rényi entropy. *IEEE Trans. Inform. Theory*, 61(2):708–714, February 2015.

[6] S. G. Bobkov, G. P. Chistyakov, and F. Götze. Rényi divergence and the central limit theorem. *Ann. Probab.*, 47(1):270–323, 2019.

[7] S. G. Bobkov and A. Marsiglietti. Variants of the entropy power inequality. *IEEE Transactions on Information Theory*, 63(12):7747–7752, 2017.

[8] S. G. Bobkov and A. Marsiglietti. Asymptotic behavior of Rényi entropy in the central limit theorem. *Preprint, arXiv:1802.10212*, 2018.

[9] S. Brazitikos, A. Giannopoulos, P. Valettas, and B.-H. Vritsiou. Geometry of isotropic convex bodies. *AMS-Mathematical Surveys and Monographs 196*, 2014.

[10] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. J. Wiley, New York, 1991.

[11] B. V. Gnedenko and 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 Co., Reading, Mass.-London-Don Mills., Ont., 1968.

[12] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[13] O. Johnson. *Information theory and the central limit theorem*. Imperial College Press, London, 2004.

[14] J. Li. Rényi entropy power inequality and a reverse. *Studia Math*, 242:303 – 319, 2018.

[15] J. Li, A. Marsiglietti, and J. Melbourne. Further investigations of Rényi entropy power inequalities and an entropic characterization of s-concave densities. *Preprint,* arXiv:1901.10616 [math.PR], 2019.

[16] J. Li, A. Marsiglietti, and J. Melbourne. Rényi entropy power inequalities for s-concave densities. In *IEEE International Symposium on Information Theory*, Paris, France, 2019.

[17] M. Madiman and A. R. Barron. Generalized entropy power inequalities and monotonicity properties of information. *IEEE Trans. Inform. Theory*, 53(7):2317–2329, July 2007.

[18] M. Madiman, J. Melbourne, and P. Xu. Forward and reverse entropy power inequalities in convex geometry. *Convexity and Concentration*, pages 427–485, 2017.

[19] M. Madiman, J. Melbourne, and P. Xu. Rogozin's convolution inequality for locally compact groups. *Preprint, arXiv:1705.00642*, 2017.

[20] A. Marsiglietti and J. Melbourne. A Rényi entropy power inequality for log-concave vectors and parameters in $[0,1]$. In *Proceedings IEEE International Symposium on Information Theory*, Vail, USA, 2018.

[21] A. Marsiglietti and J. Melbourne. On the entropy power inequality for the Rényi entropy of order [0, 1]. *IEEE Transactions on Information Theory*, 65(3):1387–1396, 2019.

[22] A. Prékopa. On logarithmic concave measures and functions. *Acta Sci. Math. (Szeged)*, 34:335–343, 1973.

[23] E. Ram and I. Sason. On Rényi entropy power inequalities. *IEEE Transactions on Information Theory*, 62(12):6800–6815, 2016.

[24] O. Rioul. Rényi entropy power inequalities via normal transport and rotation. *Entropy*, 20(9):641, 2018.

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

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