Rényi Entropy and Variance Comparison for Symmetric Log-Concave Random Variables

Maciej Białobrzeski and Piotr Nayar

Abstract—We show that for any $\alpha > 0$ the Rényi entropy of order $\alpha$ is minimized, among all symmetric log-concave random variables with fixed variance, either for a uniform distribution or for a two sided exponential distribution. The first case occurs for $\alpha \in (0, \alpha^*)$ and the second case for $\alpha \in [\alpha^*, \infty)$, where $\alpha^*$ satisfies the equation $2 \log \alpha^* = (\alpha^* - 1) \log 6$, that is $\alpha^* \approx 1.241$. We deduce that the one-sided exponential distribution minimizes Rényi entropy of order $\alpha \geq 2$ among all log-concave random variables with fixed variance.

Index Terms—Rényi entropy, log-concave random variables, relative $\alpha$-entropy, entropy power inequality, variance.

I. INTRODUCTION

For a random variable $X$ with density $f$ its Rényi entropy of order $\alpha \in (0, \infty) \setminus \{1\}$ is defined as

$$h_\alpha(X) = h_\alpha(f) = \frac{1}{1-\alpha} \log \left( \int f^\alpha(x) \, dx \right),$$

assuming that the integral converges, see [25]. If $\alpha \to 1$ one recovers the usual Shannon differential entropy $h(f) = h_1(f) = -\int f \ln f$. Also, by taking limits one can define $h_0(f) = \log \|f\|_\infty$, where $\|f\|_\infty$ stands for the support of $f$ and $h_\infty(f) = -\log \|f\|_\infty$, there $\|f\|_\infty$ is the essential supremum of $f$.

It is a well known fact that for any random variable one has

$$h(X) \leq \frac{1}{2} \log \text{Var}(X) + \frac{1}{2} \log (2\pi e)$$

with equality only for Gaussian random variables, see e.g. Theorem 8.6.5 in [11]. The problem of maximizing Rényi entropy under fixed variance has been considered independently by Costa et al. in [10] and by Lutwak et al. in [17], where the authors showed, in particular, that for $\alpha \in (\frac{1}{3}, \infty) \setminus \{1\}$ the maximizer is of the form

$$f(x) = c_0(1 + (1 - \alpha)(c_1 x^2)^{\frac{1}{1-\alpha}}),$$

which will be called the generalized Gaussian density. Any density satisfying $f(x) \sim x^{-3}(\log x)^{-2}$ shows that for $\alpha \leq \frac{1}{3}$ the supremum of $h_\alpha$ under fixed variance is infinite.

Manuscript received 12 January 2022; revised 3 September 2022; accepted 24 January 2023. Date of publication 22 February 2023; date of current version 19 May 2023. The work of Piotr Nayar was supported by the National Science Centre, Poland, under Grant 2018/31/D/ST1/01355. (Corresponding author: Piotr Nayar.)

The authors are with the Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, 02-097 Warsaw, Poland (e-mail: nayar@impan.edu.pl).

Communicated by A. Tchamkerten, Associate Editor for Shannon Theory. Digital Object Identifier 10.1109/TIT.2023.3243956

One may also ask for reverse bounds. However, the infimum of the functional $h_\alpha$ under fixed variance is $-\infty$ as can be seen by considering $f_n(x) = \frac{n}{2} 1_{[1,1+n^{-1}]}(x)$ for which the variance stays bounded whereas $h_\alpha(f_n) \to -\infty$ as $n \to \infty$. Therefore, it is natural to restrict the problem to a certain natural class of densities, in which the Rényi entropy remains lower bounded in terms of the variance. In this context it is natural to consider the class of log-concave densities, namely densities having the form $f = e^{-V}$, where $V : \mathbb{R} \to (-\infty, \infty]$ is convex. In [20] it was proved that for any symmetric log-concave random variable one has

$$h(X) \geq \frac{1}{2} \log \text{Var}(X) + \frac{1}{2} \log 12$$

with equality if and only if $X$ is a uniform random variable.

In the present article we shall extend this result to general Rényi entropy. Namely, we shall prove the following theorem.

Theorem 1: Let $X$ be a symmetric log-concave random variable in $\mathbb{R}$ and $\alpha > 0$, $\alpha \neq 1$. Define $\alpha^*$ to be the unique solution to the equation $\frac{1}{\alpha-1} \log \alpha = \frac{1}{2} \log 6$ ($\alpha^* \approx 1.241$).

Then

$$h_\alpha(X) \geq \frac{1}{2} \log \text{Var}(X) + \frac{1}{2} \log 12$$

for $\alpha \leq \alpha^*$.

and

$$h_\alpha(X) \geq \frac{1}{2} \log \text{Var}(X) + \frac{1}{2} \log 2 + \frac{\log \alpha}{\alpha - 1}$$

for $\alpha \geq \alpha^*$.

For $\alpha < \alpha^*$ equality holds if and only if $X$ is uniform random variable on a symmetric interval, while for $\alpha > \alpha^*$ the bound is attained only for two-sided exponential distribution. When $\alpha = \alpha^*$, two previously mentioned densities are the only cases of equality.

The above theorem for $\alpha < 1$ trivially follows from the case $\alpha = 1$ as already observed in [20] (see Theorem 5 therein). This is due to the monotonicity of Rényi entropy in $\alpha$. As we can see the case $\alpha \in [1, \alpha^*)$ of Theorem 1 is a strengthening of the main result of [20], as in this case $h_\alpha(X) \leq h(X)$ and the right hand sides are the same.

It turns out that Theorem 1 allows to deal with the non-symmetric case in the range $\alpha \geq 2$. The following corollary of our main theorem has been kindly communicated to us by Jiange Li.

Corollary 2: Let $X$ be a log-concave random variable in $\mathbb{R}$ and let $\alpha \geq 2$. Then

$$h_\alpha(X) \geq \frac{1}{2} \log \text{Var}(X) + \frac{\log \alpha}{\alpha - 1}$$

with equality for one-sided exponential random variable.
To prove it we shall use Theorem 6.1 from [21]: for any iid log-concave random variables and \( \alpha \geq 2 \) one has \( h_\alpha(X-Y) \leq h_\alpha(X) + \log 2 \). Since \( X-Y \) is log-concave and symmetric, we obtain

\[
\begin{align*}
    h_\alpha(X) & \geq h_\alpha(X-Y) - \log 2 \\
    & \geq \frac{1}{2} \log \text{Var}(X-Y) - \frac{1}{2} \log 2 + \frac{\log \alpha}{\alpha - 1} \\
    & = \frac{1}{2} \log \text{Var}(X) + \frac{\log \alpha}{\alpha - 1}.
\end{align*}
\]

We remark that the problem of minimizing the Rényi entropy of order \( \alpha \in (0,2) \) under fixed variance is open in the class of arbitrary log-concave densities (not necessarily symmetric).

Let us also mention that it is natural to ask for a multi-dimensional analogue of Theorem 1. This is wide open even for \( \alpha = 1 \). In this case a relevant affine invariant inequality would be of the form \( h(X) \geq \frac{1}{2} \log(\det(\text{cov}(X))) + c(n) \), where \( X \) is a log-concave random vector in \( \mathbb{R}^n \) and we ask \( c(n) \) to be optimal. However, as explained in [5] (see also the discussion in Section III of [20]), proving such an optimal bound would imply solving the famous hyperplane conjecture due to Bourgain (see [9]), saying that every compact convex set of volume 1 admits a hyperplane section of volume at least a universal constant. A more realistic goal would be to find the best constant \( c(n) \) in the case of continuous rotation invariant log-concave measures. Since every such measure has density of the form \( e^{-V(|z|)} \), where \( V \) is convex and non-decreasing, one can again apply the techniques used in [20] and in the present article. However, this leads to an even more complicated technical inequality than the one proved here.

This article is organized as follows. In Section II we reduce Theorem 1 to the case \( \alpha = \alpha^* \). In Section III we further simplify the problem by reducing it to simple functions via the concept of degrees of freedom. Section IV contains the proof for these simple functions. In the last section we derive two applications of our main result.

II. REDUCTION TO THE CASE \( \alpha = \alpha^* \)

The following lemma is well known. We present its proof for completeness. The proof of point (ii) is taken from [13]. As pointed out by the authors, it can also be derived from Theorem 2 in [8] or from Theorem VII.2 in [5].

**Lemma 3:** Suppose \( f \) is a probability density in \( \mathbb{R}^n \).

(i) The function \( p \mapsto \int f^p \) is log-convex on \((0,\infty)\).

(ii) If \( f \) is log-concave then the function \( p \mapsto p^n \int f^p \) is log-convex on \((0,\infty)\).

**Proof:** (i) This is a simple consequence of Hölder’s inequality.

(ii) Let \( \psi(p) = p^n \int f^p(x)dx \). The function \( f \) can be written as \( f = e^{-V} \), where \( V : \mathbb{R}^n \to (-\infty, +\infty] \) is convex. Changing variables we get \( \psi(p) = \int e^{pV(\frac{x}{p})}dx \). For any convex \( V \) the so-called perspective function \( W(z, p) = pV(\frac{z}{p}) \) is convex on \( \mathbb{R}^n \times (0, \infty) \). Indeed, for \( \lambda \in [0, 1] \), \( p_1, p_2 > 0 \) and \( z_1, z_2 \in \mathbb{R}^n \) we have

\[
\begin{align*}
    W(\lambda z_1 + (1-\lambda) z_2, \lambda p_1 + (1-\lambda)p_2) & = (\lambda p_1 + (1-\lambda)p_2)V\left(\frac{\lambda p_1 z_1 + (1-\lambda)p_2 z_2}{\lambda p_1 + (1-\lambda)p_2}\right) \\
    & \leq \lambda p_1 V\left(\frac{z_1}{p_1}\right) + (1-\lambda)p_2 V\left(\frac{z_2}{p_2}\right) \\
    & = \lambda W(z_1, p_1) + (1-\lambda)W(z_2, p_2).
\end{align*}
\]

Since \( \psi(p) = \int e^{-W(z,p)}dz \), the assertion follows from the Prekopa’s theorem from [23] saying that a marginal of a log-concave function is again log-concave.

**Remark:** The use of the term perspective function appeared in [14], however the convexity of this function was known much earlier (see e.g. page 35 in the book by Rockafellar [26]).

The next corollary is a simple consequence of Lemma 3. The right inequality of this corollary appeared in [13], whereas the left inequality is classical.

**Corollary 4:** Let \( f \) be a log-concave probability density in \( \mathbb{R}^n \). Then for any \( p \geq q > 0 \) we have

\[
0 \leq h_q(f) - h_p(f) \leq \frac{\log q}{q-1} - n \frac{\log p}{p-1}.
\]

In fact, the first inequality is valid without the log-concavity assumption.

**Proof:** To prove the first inequality we observe that due to Lemma 3 the function defined by \( \phi_1(p) = (1-p)h_p(f) \) is convex. From the monotonicity of slopes of \( \phi_1 \) we get that \( \phi_1(q) - \phi_1(1) \geq \phi_2(q) - \phi_2(1) \), which together with the fact that \( \phi_1(1) = 0 \) gives \( h_q(f) \leq h_p(f) \).

Similarly, to prove the right inequality we note that \( \phi_2(p) = n \log p + (1-p)h_p(f) \) is concave with \( \phi_2(1) = 0 \). Thus, \( \phi_2(p) - \phi_2(1) \leq \phi_2(q) - \phi_2(1) \), which gives \( \frac{n \log p}{p-1} - h_p(f) \leq \frac{n \log q}{q-1} - h_q(f) \), which finishes the proof.

Having Corollary 4 we can easily reduce Theorem 1 to the case \( \alpha = \alpha^* \). Indeed, the case \( \alpha < \alpha^* \) follows from the left inequality of Corollary 4 (\( h_p \) is non-increasing in \( p \)). The case \( \alpha > \alpha^* \) is a consequence of the right inequality of the above corollary, according to which the quantity \( h_\alpha(X) - \frac{\log \alpha}{\alpha - 1} \) is non-decreasing in \( \alpha \).

III. REDUCTION TO SIMPLE FUNCTIONS VIA DEGREES OF FREEDOM

The content of this section is a rather straightforward adaptation of the method from [20]. Therefore, we shall only sketch the arguments.

By a standard approximation argument it is enough to prove our inequality for functions from the set \( \mathcal{F}_L \) of all continuous even log-concave probability densities supported on \([-L, L] \). Thus, it suffices to show that

\[
\inf \{ h_\alpha^*(f) : f \in A \} \geq \log \sigma + \frac{1}{2} \log 2 + \frac{\log \alpha^*}{\alpha^* - 1},
\]

where \( A = \{ f \in \mathcal{F}_L : \text{Var}(f) = \sigma^2 \} \). We shall show that \( \inf_{f \in A} h_\alpha^*(f) \) is attained on \( A \). Equivalently, since \( \alpha^* > 1 \) it suffices to show that \( M = \sup_{f \in A} \int f^\alpha \) is attained on \( A \). We first argue that this supremum is finite. This follows from
the estimate \( \int f^{\alpha^*} \leq 2 L f(0)^{\alpha^*} \) and from the inequality \( f(0) \leq \frac{1}{\sqrt{2} \var(f)} = \frac{1}{\sqrt{2} \sigma} \), see Lemma 1 in [20]. Next, let \((f_n)_{n \geq 1}\) be a sequence of functions from \(A\) such that \( \int f_n^{\alpha^*} \to M \). According to Lemma 2 from [20], by passing to a subsequence one can assume that \( f_n \to f \) pointwise, where \( f \) is some function from \( A \). Since \( f_n \leq f_n(0) \leq \frac{1}{\sqrt{2} \sigma} \), by the Lebesgue dominated convergence theorem we get that \( \int f^{\alpha^*} = \int f^{\alpha^*} = M \) and therefore the supremum is attained on \( A \).

Now, we say that \( f \in A \) is an extremal point in \( A \) if \( f \) cannot be written as a convex combination of two different functions from \( A \), that is, if \( f = \lambda f_1 + (1 - \lambda) f_2 \) for some \( \lambda \in (0, 1) \) and \( f_1, f_2 \in A \), then necessarily \( f_1 = f_2 \). It is easy to observe that if \( f \) is not extremal, then it cannot be a maximizer of \( \int f^{\alpha^*} \) on \( A \). Indeed, if \( f = \lambda f_1 + (1 - \lambda) f_2 \) for some \( \lambda \in (0, 1) \) and \( f_1, f_2 \in A \) with \( f_1 \neq f_2 \), then the strict convexity of \( x \to x^{\alpha^*} \) implies

\[
\int f^{\alpha^*} = \int (\lambda f_1 + (1 - \lambda) f_2)^{\alpha^*} < \lambda \int f_1^{\alpha^*} + (1 - \lambda) \int f_2^{\alpha^*} \leq M.
\]

This shows that in order to prove (1) it suffices to consider only the functions \( f \) being extremal points of \( A \). Finally, according to Steps III and IV of the proof of Theorem 1 from [20] these extremal points are of the form

\[
f(x) = c \mathbf{1}_{[0,a]}(|x|) + e^{-\gamma(|x| - \alpha)} \mathbf{1}_{[a,\alpha + b]}(|x|), \quad a, b, \gamma \geq 0,
\]

with \( a + b \leq L, \ c > 0, \ a, b, \gamma \geq 0 \), where it is also assumed that \( \int f = 1 \).

IV. PROOF FOR THE CASE \( \alpha = \alpha^* \)

Due to the previous section, we can restrict ourselves to probability densities \( f \) of the form

\[
f(x) = c \mathbf{1}_{[0,a]}(|x|) + e^{-\gamma(|x| - \alpha)} \mathbf{1}_{[a,\alpha + b]}(|x|), \quad a, b, \gamma \geq 0.
\]

The inequality is invariant under scaling \( f(x) \mapsto \lambda f(\lambda x) \) for any positive \( \lambda \), so we can assume that \( \gamma = 1 \) (note that in the case \( \gamma = 0 \) we get equality). We have

\[
\int f^{\alpha^*} = \int c \mathbf{1}_{[0,a]}(|x|) + e^{-\alpha x} \mathbf{1}_{[a,\alpha + b]}(|x|) = 2 e^{\alpha} \left( a + 1 - e^{-ab} \right)
\]

and thus

\[
h_\alpha(f) = \frac{1}{1 - \alpha} \log \int f^{\alpha^*} = \frac{1}{1 - \alpha} \log \left( 2 e^{\alpha} \left( a + 1 - e^{-ab} \right) \right).
\]

Moreover,

\[
\var(f) = 2 c \int x^2 \mathbf{1}_{[0,a]}(x) dx + 2 c \int (x + a) e^{-x} \mathbf{1}_{[0,a]}(x) dx = 2 c \left( \frac{a^3}{3} + \int_0^b (x + a)^2 e^{-x} dx \right).
\]

so our inequality can be rewritten as

\[
\frac{1}{1 - \alpha^*} \log \left( 2 e^{\alpha^*} \left( a + 1 - e^{-\alpha^* b} \right) \right) + \log \alpha^* \geq \frac{1}{2} \log \left( 2 e \left( \frac{a^3}{3} + \int_0^b (x + a)^2 e^{-x} dx \right) \right) + \frac{1}{2} \log 2,
\]

which is

\[
\frac{1}{1 - \alpha^*} \log \left( 2 e^{\alpha^*} \left( a \alpha^* + 1 - e^{-\alpha^* b} \right) \right) \geq \frac{1}{2} \log \left( 2 e \left( \frac{a^3}{3} + \int_0^b (x + a)^2 e^{-x} dx \right) \right) + \frac{1}{2} \log 2.
\]

The constraint \( \int f_g = 1 \) gives \( c = \frac{1}{2} (a + 1 - e^{-b})^{-1} \). After multiplying both sides by 2, exponentiating both sides and plugging the expression for \( c \) in, we get the equivalent form of the inequality, \( G(a, b, \alpha^*) \geq 0 \), where

\[
G(a, b, \alpha) = 2 (a + 1 - e^{-ab}) \left( a + 1 - e^{-b} \right)^{\frac{1 - \alpha}{1 - \alpha^*}} - \left( \frac{a^3}{3} + \int_0^b (x + a)^2 e^{-x} dx \right).
\]

We will also write \( G(a, b) = G(a, b, \alpha^*) \).

To finish the proof we shall need the following lemma.

**Lemma 5**: The following holds:

(a) \( \frac{\partial^3}{\partial a^3} G(a, b) \geq 0 \) holds for every \( a, b \geq 0 \),

(b) \( \lim_{a \to \infty} \frac{\partial^3}{\partial a^3} G(a, b) = 0 \) for every \( b \geq 0 \),

(c) \( \lim_{a \to \infty} \frac{\partial^2}{\partial a^2} G(a, b) \geq 0 \) for every \( b \geq 0 \),

(d) \( \frac{\partial}{\partial a} G(a, b) \bigg|_{a=\infty} \geq 0 \) for every \( b \geq 0 \),

(e) \( G(0, b) \geq 0 \) for every \( b \geq 0 \).

With these claims at hand it is easy to conclude the proof. Indeed, one easily gets, one by one,

\[
\frac{\partial^3 G}{\partial a^3} \leq 0, \quad \frac{\partial^2 G}{\partial a^2} \geq 0, \quad \frac{\partial G}{\partial a} \geq 0, \quad G \geq 0
\]

for \( G = G(a, b) \) with fixed \( b \geq 0 \).

The proof of points (d) and (e) relies on the following simple lemma.

**Lemma 6**: Let \( f(x) = \sum_{n=0}^{\infty} a_n x^n \), where the series is convergent for every nonnegative \( x \). If there exists a nonnegative integer \( N \) such that \( a_n \geq 0 \) for \( n < N \) and \( a_n \leq 0 \) for \( n \geq N \), then \( f \) changes sign on \( (0, \infty) \) at most once. Moreover, if at least one coefficient \( a_n \) is positive and at least one negative, then there exists \( x_0 \) such that \( f(x) > 0 \) on \( [0, x_0) \) and \( f(x) < 0 \) on \( (x_0, \infty) \).

**Proof**: Clearly the function \( f(x) x^{-N} \) is nonincreasing on \( (0, \infty) \), so the first claim follows. To prove the second part
we observe that for small $x$ the function $f$ must be strictly positive and $f(x)x^{-N}$ is strictly decreasing on $(0, \infty)$. □

With this preparation we are ready to prove Lemma 5.

Proof of Lemma 5:

(a) This point is the crucial step of the proof. We will use the following observation.

Lemma 7: Let $f(x) = (x+a)^\gamma(x+b)^\delta$, where $a, b \geq 0$ and $\gamma + \delta = m \in \mathbb{N}$. Then $(m+1)$-th derivative of $f$ has fixed sign on $(0, \infty)$. The sign of $f$ is equal to $\text{sgn}(\gamma(\gamma-1)\cdots(\gamma-m))\text{sgn}((b-a)^{m+1})$.

Proof: We can write out

$$f^{(m+1)}(x) = (x+a)^{\gamma-m-1}(x+b)^{\delta-m-1} \times \left(\sum_{k=0}^{m+1} \binom{m+1}{k} c_k(x+a)^k(x+b)^{m+1-k}\right),$$

where

$$c_k = \gamma(\gamma-1)\cdots(\gamma-m+k)\delta(\delta-1)\cdots(\delta-k+1)$$

and since $\delta = m - \gamma$ we get

$$c_k = \gamma\cdots(\gamma-m+k)(m-\gamma)\cdots(m-\gamma-k+1) = (-1)^k\gamma\cdots(\gamma-m+k)(\gamma-m)\cdots(\gamma-m+k+1)
= (-1)^k C,$$

where $C = \gamma(\gamma-1)\cdots(\gamma-m)$ and does not depend on $k$.

Thus $f^{(m+1)}(x)$ simplifies to

$$C(x+a)^{\gamma-m-1}(x+b)^{\delta-m-1} \times \sum_{k=0}^{m+1} \binom{m+1}{k} (-1)^k (x+a)^k(x+b)^{m+1-k}
= C(x+a)^{\gamma-m-1}(x+b)^{\delta-m-1}(b-a)^{m+1}.$$

Since $(x+a)^{\gamma-m-1}(x+b)^{\delta-m-1} > 0$, the thesis follows. □

Now we apply Lemma 7 to (2). The expression

$$a^3 + \int_0^b (x+a)^2e^{-x}dx$$

is a polynomial of degree 3 in $a$ (which follows immediately by differentiation under the integral sign), so its fourth derivative vanishes. The remaining term is

$$2(\alpha^*)^{\frac{2-\gamma}{1-\alpha^*}}(a+\alpha^*)^{-1}e^{-\alpha^*b}(\alpha^*)^{-1}\frac{2}{1-\alpha^*}(a+1-e^{-b})^{-1}\frac{1-3a^*}{1-\alpha^*}.$$

The sum of the exponents is 3, so we can take $m = 3,\ a' = (\alpha^*)^{-1}e^{-\alpha^*b}(\alpha^*)^{-1},\ b' = 1 - e^{-b},\ \gamma = \frac{2}{1-\alpha^*}$ and $\delta = \frac{1-3a^*}{1-\alpha^*}$ in Lemma 7. Since $\gamma < 0$, the sign is equal to $(-1)^4\text{sgn}(b-a)^4 = 1$.

(b-c) Let us denote $\tau_a = 1 - e^{-ab}$ and $\tau = 1 - e^{-b}$. We have

$$G(a, b, \alpha) = 2(aa + \alpha^*)^{\frac{2}{1-\alpha^*}}(a + \tau)^\frac{2}{1-\alpha^*} - \left(\frac{a^3}{3} + \int_0^b (x+a)^2e^{-x}dx\right)\frac{a^2}{2(\alpha^*)^{\frac{2-\gamma}{1-\alpha^*}}(a+\alpha^*)^{-1}e^{-\alpha^*b}(\alpha^*)^{-1}\frac{2}{1-\alpha^*}(a+1-e^{-b})^{-1}\frac{1-3a^*}{1-\alpha^*}}$$

$$= 2\ a^3\ a^{\frac{2}{1-\alpha^*}}\left(1 + \frac{\tau_a}{a\alpha^*}\right)^\frac{2}{1-\alpha^*}\left(1 + \frac{\tau}{a}\right)^{\frac{2a}{3a-1}} - \left(\frac{a^3}{3} + \int_0^b (x+a)^2e^{-x}dx\right).$$

We now Taylor expand the term $\left(1 + \frac{\tau_a}{a\alpha^*}\right)^\frac{2}{1-\alpha^*}\left(1 + \frac{\tau}{a}\right)^{\frac{2a}{3a-1}}$ in $a^{-1}$ to get that as $a \to \infty$

$$G(a, b, \alpha) = t_3\alpha^3 + t_2\alpha^2 + t_1\alpha + t_0 + t_{-1}\alpha^{-1} + t_{-2}\alpha^{-2} + \ldots,$$

where

$$t_3(\alpha) = 2\alpha^{\frac{2}{3}} - \frac{1}{3}$$

and

$$t_2(\alpha) = 2\alpha^{\frac{2}{3}} \left(\frac{2\tau_a}{\alpha(1-\alpha)} + \frac{3\alpha - 1}{\alpha - 1}\right) - \tau$$

$$= \frac{2\alpha^{\frac{2}{3}}}{\alpha - 1} (\alpha(3\alpha - 1)\tau - 2\tau_a) - \tau.$$

We have $t_3(\alpha^*) = 0$. Therefore, denoting $t_k^* = t_k(\alpha^*)$, we get

$$G(a, b, \alpha) = 2\ t_2^*\alpha^2 + t_1^*\alpha + t_0 + t_{-1}\alpha^{-1} + t_{-2}\alpha^{-2} + \ldots$$

Thus $\frac{\partial^2 G(a, b, \alpha)}{\partial \alpha^2} = 2\ t_2^* + 2\ t_1^*\alpha - 3\ t_2^*\alpha^2 + \ldots$ and thus $\lim_{\alpha \to \infty} \frac{\partial^2 G(a, b, \alpha)}{\partial \alpha^2} = 2t_2^*$. Using the equality $2(\alpha^*)^{-\frac{2-\gamma}{1-\alpha^*}} = \frac{1}{1-\alpha^*}$ we get

$$\lim_{\alpha \to \infty} \frac{\partial^2 G(a, b, \alpha)}{\partial \alpha^2} = \frac{2}{3(\alpha^*)} (\alpha^*(3\alpha^* - 1)\tau - 2\tau_a) - 2\tau$$

$$= \frac{4}{3(\alpha^*)} (\tau\alpha^* - \tau_a).$$

This expression is nonnegative for $b \geq 0$, since the function $h_3(x) = 1 - e^{-x}$ is concave, so we have $\frac{1}{1-\alpha^*} > \frac{h_3(b)}{\alpha^*} = \frac{1-e^{-\alpha^*b}}{\alpha^*}$ as $\alpha^* > 1$ (monotonicity of slopes).

(e) To illustrate our method, before proceeding with the proof of (d) we shall prove (e), as the idea of the proof of (d) is similar, but the details are more complicated. Our goal is to show the inequality

$$(1 - e^{-ab})^{\frac{2}{1-\alpha^*}}(1 - e^{-b})^{\frac{1-3a^*}{1-\alpha^*}} \geq 1 - \frac{b^2 + 2b + 2}{2} e^{-b}.$$ (3)

after taking the logarithm of both sides our inequality reduces to nonnegativity of

$$\phi(b) = \frac{2}{1-\alpha^*} \log(1 - e^{-\alpha^*b}) + \frac{1 - 3a^*}{1-\alpha^*} \log(1 - e^{-b})$$

$$- \log \left(1 - \frac{b^2 + 2b + 2}{2} e^{-b}\right).$$

We have

$$\phi'(b) = \frac{2\alpha^*}{(1-\alpha^*)(e^{\alpha^*b} - 1)} + \frac{1 - 3a^*}{b^2}$$

$$+ \frac{b^2 + 2b - 2e^b + 2}{2}.$$ (4)

It turns out that $\phi(b)$ changes sign on $(0, \infty)$ at most once. To show that, firstly, clear out the denominators (they have fixed sign on $(0, \infty)$) to obtain the expression

$$2\alpha^*(b^2 + 2b - 2e^b + 2)(e^b - 1)$$

$$+ (1 - 3a^*)(e^{\alpha^*b} - 1)(b^2 + 2b - 2e^b + 2)$$

$$+ b^2(1 - \alpha^*)(e^b - 1)(e^{\alpha^*b} - 1).$$
Now we will apply Lemma 6 to (4). That expression can be rewritten as
\[
-4\alpha^* \left( \sum_{n=3}^{\infty} \frac{b^n}{n!} \right) \left( \sum_{n=1}^{\infty} \frac{b^n}{n!} \right) \\
+ (6\alpha^* - 2) \left( \sum_{n=1}^{\infty} \frac{(\alpha^b)^n}{n!} \right) \left( \sum_{n=3}^{\infty} \frac{b^n}{n!} \right) \\
+ b^2(1 - \alpha^*) \left( \sum_{n=1}^{\infty} \frac{b^n}{n!} \right) \left( \sum_{n=1}^{\infty} \frac{(\alpha^b)^n}{n!} \right),
\]
so the \(n\)-th coefficient \(a_n\) in the Taylor expansion is equal to
\[
a_n = (6\alpha^* - 2) \left( \sum_{j=1}^{\infty} \frac{(\alpha^*)^j j!}{(n-j)!} \right) \\
+ (1 - \alpha^*) \left( \sum_{j=1}^{\infty} \frac{\alpha^*^j}{j!(n-j)!} \right) \\
\leq \frac{1}{n!} (6\alpha^* - 2)(\alpha^* + 1)^n \\
+ \frac{1 - \alpha^*}{(n - 2)!} ((\alpha^* + 1)^{n-2} - (\alpha^*)^{n-2}) \\
\leq \frac{6}{n!} (\alpha^* + 1)^n - \frac{n(n-1)}{30n!} (\alpha^* + 1)^n + \frac{8n^2}{n^2} (\alpha^*)^n.
\]
When \(n \geq 17\), we have \(\frac{n(n-1)}{30} > 7\) and \((\frac{6}{\alpha^*})^n \geq 8^n\), so \(a_n\) is less than zero for \(n \geq 17\). It can be checked (preferably using computational software) that the rest of coefficients \(a_n\) satisfy the pattern from Lemma 6, with \(a_n = 0\) for \(n \leq 4\), \(a_n > 0\) for \(n = 5, 6, 7\) and \(a_n < 0\) for \(n > 8\).

This way we have proved that \(\phi^\prime(b)\) changes sign in exactly one point \(x_0 \in (0, \infty)\). Thus, \(\phi\) is first increasing and then decreasing. Since \(\phi(0) = 0\) and \(\lim_{b \to \infty} \phi(b) = 0\), the assertion follows.

(d) We have to show that
\[
F(b) = P(b)Q(b) + R(b),
\]
where
\[
P(b) = - (e^b - 1) (e^b - 1 - b) (\alpha^* + 1)\alpha^* \\
+ 2 (e^b - 1 - b) \alpha^* (e^{\beta b^*} - 1) \\
- b (e^b - 1) (\alpha^* - 1) (e^{\beta b^*} - 1) \\
Q(b) = e^b (1 - 3\alpha^*) + 2\alpha^* e^{\beta b^*} + (\alpha^* - 1) e^{\beta b^* + b} \\
R(b) = \alpha^* (\alpha^* - 1) (e^b - 1) (e^b - 1 - b) \left( e^{\beta b^*} - 1 \right) \\
\times \left( e^b (3\alpha^* - 1) - 2e^{\beta b^*} \right).
\]
Let us consider the Taylor series \(\sum_{n=0}^{\infty} f_n b^n\) of \(F(b)\) (it is clear that the series converges to the function everywhere). We first notice that the coefficients \(f_n\) can be derived explicitly as \(F(b)\) is a sum of functions of the form \(b^k e^{\beta b} = \sum_{n=0}^{\infty} \frac{c_n}{n!} b^{n+k}\), where \(k \geq 0\) is an integer and \(c \in \mathbb{R}\). It can be checked (again using computational software) that coefficients of this series up to order 9 are nonnegative and coefficients of order greater than 9, but lesser than 30 are negative. In order to do this one can e.g. derive the \(n\)th coefficient symbolically using the command \(\text{SeriesCoefficient}[F(b,\alpha), \{b, 0, n\}]\) in Mathematica. This gives a long symbolic expression in \(n\) and \(\alpha^*\), which we prefer not to write here. It is then straightforward to evaluate this expression numerically for \(n \in \{0, \ldots, 29\}\) (the symbolic value of \(\alpha^*\) that can be evaluated numerically with high precision is obtained by using the command \(\text{Solve}[2 \log[x] == (x - 1) \log[6] \&\& x > 1]\)). Now we will show negativity of coefficients \(f_n\) of order at least 30 (our bound will be very crude, so it would not work, if we replaced 30 with lower number). Firstly we note that \(Q(b)\) has \(n\)-th Taylor coefficient equal to
\[
\frac{1 - 3\alpha^* + 2(\alpha^*)^{n+1} + (\alpha^* - 1)(\alpha^* + 1)^n}{n!} \geq \frac{1 - 3\alpha^* + 2\alpha^* + \alpha^* - 1}{n!} = 0,
\]
so all its coefficients are nonnegative. If \(P(b) = \sum p_n b^n\) and \(Q(b) = \sum q_n b^n\) then the \(n\)th Taylor coefficient of \(PQ\) is given by \(\sum_{k=0}^{n} p_k q_{n-k}\). Since \(q_{k} \geq 0\), if we replace the function \(P\) with a function \(\tilde{P}\) = \(\sum \tilde{p}_k b^k\) having bigger coefficients \(\tilde{p}_k \geq p_k\), then the coefficients of \(\tilde{P}Q\) will be bigger than those of \(PQ\). We have \(P = P_1 + P_2 + P_3\), where
\[
P_1(b) = - (e^b - 1) (e^b - 1 - b) (\alpha^* + 1)\alpha^* \\
P_2(b) = 2 (e^b - 1 - b) \alpha^* (e^{\beta b^*} - 1) \\
P_3(b) = -b (e^b - 1) (\alpha^* - 1) (e^{\beta b^*} - 1).
\]
We shall form \(\tilde{P} = \tilde{P}_1 + \tilde{P}_2 + \tilde{P}_3\) by modifying each of the functions \(P_i\). As \(P_1\) has nonpositive Taylor coefficients, we can put \(\tilde{P}_1 = 0\). If \(g\) does not have smaller coefficients than \(f\), we shall denote this by \(f \prec g\). Note that the coefficients of
$P_2$ are nonnegative and $2\alpha^* < 5/2$. We therefore have

$$P_2(b) < \frac{5}{2} \left( e^b - 1 - b \right) \left( e^{b\alpha^*} - 1 \right)$$

$$< \frac{5}{2} \left( e^b - 1 \right) \left( e^{b\alpha^*} - 1 \right) =: \tilde{P}_2(b).$$

Moreover, $\alpha^* - 1 > 1/5$ and thus

$$P_3(b) < -\frac{b}{5} \left( e^b - 1 \right) \left( e^{b\alpha^*} - 1 \right) =: \tilde{P}_3(b).$$

Now we want to show the negativity of coefficients of order at least 30 for $PQ+R$, which is equal to $(e^b-1)(e^{b\alpha^*}-1)W(b)$, where

$$W(b) = \left( \frac{5}{2} - \frac{b}{5} \right) \left( e^b(1-3\alpha^*) + 2\alpha^* e^{b\alpha^*} \right)$$

$$+ \alpha^*(\alpha^* - 1) e^{b\alpha^*} b +$$

$$\alpha^*(\alpha^* - 1)(e^b - b - 1) ((3\alpha^* - 1)e^b - 2e^{b\alpha^*}).$$

The function $W(b)$ has $n$-th Taylor coefficient $w_n$ equal to zero for $n \in \{0,1\}$, while for $n \geq 2$ it is

$$w_n = \frac{5(1-3\alpha^*)}{2n!} + \frac{3\alpha^* - 1}{5(n-1)!} + \frac{5\alpha^*(n+1)}{n!} - \frac{2(\alpha^*)^n}{5(n-1)!}$$

$$+ \frac{5(\alpha^* - 1)(\alpha^* + 1)^n}{2n!} - \frac{(\alpha^* - 1)(\alpha^* + 1)^{n-1}}{n!}$$

$$+ \frac{\alpha^*(\alpha^* - 1)3\alpha^* - 1}{n!} - \frac{2\alpha^*(\alpha^* - 1)}{n!} (((\alpha^* + 1)^n - (\alpha^*)^n - (\alpha^*)^{n-1}).$$

Again, this can be derived by opening the brackets and writing the function as a sum of functions of the form $b^e e^b$ or by using software for symbolic computation.

Upper bounding the first and the fourth term by zero, using the bound

$$\alpha^*(\alpha^* - 1)(3\alpha^* - 1) - \frac{2n}{n!} \leq \frac{2n}{n!}$$

and

$$\frac{2\alpha^*(\alpha^* - 1)}{n!} ((\alpha^*)^n + n(\alpha^*)^{n-1}) \leq \frac{(n+1)(\alpha^*)^n}{n!}$$

together with

$$-\frac{(\alpha^* - 1)(\alpha^* + 1)^n}{5(n-1)!} \leq -\frac{\frac{4}{5}n}{10n!} (\alpha^* - 1)(\alpha^* + 1)^n,$$

we get the following upper bound for $w_n$ for $n \geq 2$

$$w_n \leq \frac{(3\alpha^* - 1)}{5(n-1)!} + \frac{5\alpha^*(n+1)}{2n!} + \frac{(n+1)(\alpha^*)^n}{n!}$$

$$+ \frac{(\alpha^* + 1)^n(\alpha^* - 1)(25 - 20\alpha^* - 4n/5)}{10n!}$$

$$\leq \frac{(n+8)(\alpha^*)^n}{n!} + \frac{n+2}{n!} + \frac{(\alpha^* + 1)(1-3n)}{200n!},$$

since $\frac{1}{10}(\alpha^* - 1)(25 - 20\alpha^*) \leq \frac{1}{200}$ and $\frac{1}{200}(\alpha^* - 1) \geq \frac{3}{200}$. This bound works for $n \in \{0,1\}$, too. We have

$$(e^b - 1)(e^{b\alpha^*} - 1) = \sum_{n=2}^{\infty} b^n (\alpha^* + 1)^n - (\alpha^*)^n - 1.$$
By using the equality $(\alpha^*)^{-\frac{2}{\alpha}} = \frac{1}{6}$ we see that the constant term vanishes. In fact
\[
\phi_1(b) - \phi_2(b) = \left( \frac{1}{3} - (\alpha^*)^{-\frac{2}{\alpha}} \right) b + O(b^2) = \frac{1}{6} b + O(b^2).
\]
\[\square\]

V. APPLICATIONS

A. Relative $\alpha$-Entropy

Recall that if $f_X$ denotes the density of a random variable $X$ then the relative $\alpha$-entropy studied by Ashok Kumar and Sundaresan in [1] is defined as
\[
I_\alpha(X||Y) = \frac{\alpha}{1 - \alpha} \log \left( \int \frac{f_X}{\|f_X\|_\alpha} \left( \frac{f_Y}{\|f_Y\|_\alpha} \right)^{\alpha - 1} \right)
\]
for $\alpha \in (0, 1) \cup (1, \infty)$, where $\|f\|_\alpha = (\int |f|^\alpha)^{1/\alpha}$. We shall derive an analogue of Corollary 5 from [20]. To this end we shall need the following fact.

Proposition 8 ([1], Corollary 13): Suppose $\alpha > 0$, $\alpha \neq 1$ and let $\mathcal{P}$ be the family of probability measures such that the mean of the function $T : \mathbb{R} \rightarrow \mathbb{R}$ under them is fixed at a particular value $t$. Let the random variable $X$ have a distribution from $\mathcal{P}$, and let $Z$ be a random variable that maximizes the Rényi entropy of order $\alpha$ over $\mathcal{P}$. Then
\[
I_\alpha(X||Z) \leq h_\alpha(Z) - h_\alpha(X).
\]
Combining Proposition 8 with Theorem 1 and using expressions for the Rényi entropy and variance of a generalized Gaussian density derived in [17], one gets the following corollary.

Corollary 9: Suppose $\alpha > 1$. Let $X$ be a symmetric log-concave random variable. Let $Z$ be the random variable having generalized Gaussian density with parameter $\alpha$ and satisfying $\text{Var}(X) = \text{Var}(Z)$. Then $I_\alpha(X||Z) \leq C(\alpha)$, where
\[
C(\alpha) = \alpha \log \left( \frac{2^{\alpha}}{3\alpha - 1} \right) + \frac{1}{2} B \left( \frac{\alpha}{2}, \frac{\alpha}{\alpha - 2} \right) - \frac{2\log 2 + \log \alpha}{\alpha - 1}.
\]
Here $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ stands for the Beta function.

B. Reverse Entropy Power Inequality

The Rényi entropy power of order $\alpha > 0$ of a random vector $X$ in $\mathbb{R}^n$ is defined as $N_\alpha(X) = \exp \left( \frac{2}{n} h_\alpha(X) \right)$. We also write $N(X)$ for $N_1(X)$. If we combine our Theorem 1 with Theorem 2 from [17], we get the following sandwich bound for $\alpha > 1$ and a symmetric log-concave random variable $X$,
\[
C_-(\alpha) \text{Var}(X) \leq N_\alpha(X) \leq C_+(\alpha) \text{Var}(X),  \tag{5}
\]
where
\[
C_-(\alpha) = \begin{cases} 
12 & \alpha \in (1, \alpha^*) \\
2\alpha^{-2} & \alpha \geq \alpha^*
\end{cases}
\]
and
\[
C_+(\alpha) = \frac{3\alpha - 1}{\alpha - 1} \left( \frac{2\alpha}{3\alpha - 1} \right)^{\frac{1}{\alpha}} B \left( \frac{1}{2}, \frac{\alpha}{\alpha - 1} \right)^2.
\]

The case of $\alpha \in \left( \frac{1}{2}, 1 \right]$ was discussed in [20]. We point out that for the upper bound the log-concavity assumption is not needed. Nevertheless, note that for $\alpha > 1$ the so called generalized Gaussian density for which the right inequality is saturated, is symmetric and log-concave.

We can now easily derive an analogue of Corollary 6 from [20] for $\alpha > 1$.

Corollary 10: Let $\alpha \geq 1$. For $X, Y$ uncorrelated, symmetric real log-concave random variables one has
\[
N_\alpha(X + Y) \leq \frac{C_+(\alpha)}{C_-(\alpha)} \left( N_\alpha(X) + N_\alpha(Y) \right).
\]

Proof of Lemma 5: We have
\[
N_\alpha(X + Y) \leq C_+(\alpha) \text{Var}(X + Y),
\]
and
\[
= C_+(\alpha)(\text{Var}(X) + \text{Var}(Y)) \leq \frac{C_+(\alpha)}{C_-(\alpha)} \left( N_\alpha(X) + N_\alpha(Y) \right).
\]

Using bounds from Corollary 2 one easily get an inequality analogous to (5) in the non-symmetric case, namely for $\alpha > 2$ and an arbitrary log-concave random variable one has
\[
\frac{1}{2} C_-(\alpha) \text{Var}(X) \leq N_\alpha(X) \leq C_+(\alpha) \text{Var}(X).
\]
It is also straightforward to get an analog of Corollary 10.

Corollary 11: Let $\alpha \geq 2$. For $X, Y$ uncorrelated real log-concave random variables one has
\[
N_\alpha(X + Y) \leq \frac{2C_+(\alpha)}{C_-(\alpha)} \left( N_\alpha(X) + N_\alpha(Y) \right).
\]

Let us point out that many other reversals of the celebrated entropy power inequality (EPI) of Shannon [27] and Stam [28] has been established. Firstly, one should point out that according to the work of Bobkov and Chistyakov [3] no reverse EPI can be formulated for general independent random variables. Indeed, there exists $X$ with finite entropy and such that $h(X + Y) = \infty$ for every independent $Y$ with finite entropy. Bobkov and Madiman in [6] showed that for any pair $X, Y$ of independent log-concave random vectors in $\mathbb{R}^n$ there exist affine entropy preserving transformations $u, v : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that
\[
N(u(X) + v(Y)) \leq C(N(X) + N(Y)),
\]
where $C$ is a universal constant. This is sometimes called the positional reverse entropy power inequality. An analogue of this result for Rényi entropy is given in [19].

In [2] Ball et al. showed that for any symmetric log-concave random vector $(X, Y)$ in $\mathbb{R}^2$ (in particular, for $X, Y$ being independent symmetric real random variables) one has
\[
N^{\frac{1}{2}}(X + Y) \leq e(N^{\frac{1}{2}}(X) + N^{\frac{1}{2}}(Y))
\]
and conjectured that the inequality holds with constant $1$ instead of $e$. The authors proved also that $N^{\kappa}(X + Y) \leq N^{\kappa}(X) + N^{\kappa}(Y)$ holds true with $\kappa = \frac{1}{110}$ in the above setting. In [19] Madiman et al. established the same inequality for arbitrary Rényi entropy power of order $\alpha \geq 1$. In fact their constant $\kappa$ depends on $\alpha$ and is always better than $\kappa = \frac{1}{110}$. In [15] Li showed that $N^{\alpha/2}(X + Y) \leq N^{\alpha/2}(X) + N^{\alpha/2}(Y)$ holds true with
communicating to them the fact that Theorem 1 combined with 
be found in the survey article [19].

[3], [4], [7], [15], [18], [24], [29]. More information on various 
EPI for two iid summands see [5], [16], [19].

by Madiman, Tkocz and the second named author. For reverse 
EPI for two iid summands see [5], [16], [19].

For results concerning the forward Rényi EPI see 
[3], [4], [7], [15], [18], [24], [29]. More information on various

ACKNOWLEDGMENT

The authors would like to thank Jiange Li for communicating to them the fact that Theorem 1 combined with Theorem 6.1 from [21] yields Corollary 2.

REFERENCES

[1] M. Ashok Kumar and R. Sundaresan, “Minimization problems based on relative α-entropy I: Forward projection,” IEEE Trans. Inf. Theory, vol. 61, no. 9, pp. 5063–5080, Jun. 2015.

[2] K. Ball, P. Nayar, and T. Tkocz, “A reverse entropy power inequality for log-concave random vectors,” Studia Math., vol. 235, pp. 17–30, Sep. 2016.

[3] S. G. Bobkov and G. P. Chistyakov, “Entropy power inequality for the Rényi entropy,” IEEE Trans. Inf. Theory, vol. 61, no. 2, pp. 708–714, Feb. 2015.

[4] S. 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), vol. 408, pp. 100–106, May 2012.

[5] S. Bobkov and M. Madiman, “The entropy per coordinate of a random vector is highly constrained under convexity conditions,” IEEE Trans. Inf. Theory, vol. 57, no. 8, pp. 4940–4954, Aug. 2011.

[6] S. Bobkov and M. Madiman, “Dimensional behaviour of entropy and information,” Comp. Rendus Mathematique, vol. 349, nos. 3–4, pp. 201–204, Feb. 2011.

[7] S. G. Bobkov and A. Marsiglietti, “Variants of the entropy power inequality,” IEEE Trans. Inf. Theory, vol. 63, no. 12, pp. 7747–7752, Dec. 2017.

[8] C. Borell, “Complements of Lyapunov’s inequality,” Mathematische Annalen, vol. 205, no. 4, pp. 323–331, Dec. 1973.

[9] J. Bourgain, “On high dimensional maximal functions associated to convex bodies,” Amer. J. Math., vol. 108, no. 6, p. 1467, Dec. 1986.

[10] J. Costa, A. Hero, and C. Vignat, “On solutions to multivariate absolute entropy problems,” in Proc. Int. Workshop Energy Minimization Methods Comput. Vis. Pattern Recognit., Jul. 2003, pp. 211–226.

[11] T. M. Cover and J. A. Thomas, Elements of Information Theory, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

[12] M. Fradelizi, J. Li, and M. Madiman, “Concentration of information content for convex measures,” Electron. J. Probab., vol. 25, pp. 1–22, Jan. 2020.

[13] M. Fradelizi, M. Madiman, and L. Wang, “Optimal concentration of information content for log-concave densities,” in High Dimensional Probability VII, C. Houdré, D. Mason, P. Reynaud-Bouret, and J. Rosinski, Eds. Basel, Switzerland: Birkhäuser, 2016.

[14] J.-B. Hiriart-Urruty and C. Lemaréchal, Convex Analysis and Minimization Algorithms, I, of Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 305. Berlin, Germany: Springer-Verlag, 1993.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.