Rényi Entropy Power Inequalities for s-concave Densities

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*—In this paper, we investigate the role of convexity in entropy power inequalities. We establish Rényi entropy power inequalities of order $r \in (0,1)$ for a large class of densities, the so-called s-concave densities. This extends recent works on Rényi entropy power inequalities.**

***Index Terms*—Rényi entropy; entropy power inequality; s-concave density.**

I. Introduction

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

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

For $r \in \{0,1,\infty\}$, the Rényi entropy can be extended continuously such that the RHS of (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

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

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

The classical entropy power inequality (henceforth, EPI) of Shannon [23] and Stam [24], states that the entropy power $N(X)$ is super-additive on the sum of independent random vectors. There has been recent success on extensions of the EPI from the Shannon differential entropy to r-Rényi entropy. In [2], [3], 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

$$N_r(X_1 + \cdots + X_n) \geq c \sum_{i=1}^n N_r(X_i). \quad (3)$$

Ram and Sason soon after gave a sharpened summation dependent constant for $r \in (1,\infty)$ in [20]. For $r=\infty$ see [16], [17]. Savaré and Toscani [22] showed that a modified Rényi entropy power was concave along solutions of some nonlinear heat equation, which generalizes Costa's concavity of entropy power [8]. Bobkov and Marsiglietti [4] proved the following variant of Rényi EPI

$$N_r(X+Y)^\alpha \geq N_r(X)^\alpha + N_r(Y)^\alpha \quad (4)$$

for $r>1$ and some exponent α only depending on r. It is clear that (4) holds for more than two summands. A refinement of the exponent α was given by Li [11].

Both (3) and (4) follow from Young's convolution inequality and the entropy comparison inequality $h_{r_1}(X) \geq h_{r_2}(X)$ for any $r_1 \leq r_2$. The latter is an immediate consequence of Jensen's inequality. Analogues of (3) and (4) for Rényi entropies of order $r \in (0,1)$ require a reverse entropy comparison inequality aforementioned. This technical issue prevents Rényi EPIs of order $r \in (0,1)$ for all random vectors. In [12], the authors show that a general Rényi EPI of the form (3) indeed fails for all $r \in (0,1)$.

Theorem 1.1 ([12]): For any $r \in (0,1)$ and $\varepsilon > 0$, there exist independent random vectors $X_1, \dots, X_n$ in $\mathbb{R}^d$, for some $d \geq 1$ and $n \geq 2$, such that

$$N_r(X_1 + \cdots + X_n) < \varepsilon \sum_{i=1}^n N_r(X_i). \quad (5)$$

However, there exists a large class of densities, the so-called s-concave densities, which satisfy a reverse entropy comparison. In this paper, we will establish Rényi EPIs of order $r \in (0,1)$ for such densities. This extends the results for log-concave densities in [18], [19].

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

$$f((1-\lambda)x + \lambda y) \geq ((1-\lambda)f(x)^s + \lambda f(y)^s)^{1/s} \quad (6)$$

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 (6) 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 an s-concave density was initiated by Borell in the seminal work [5], [6]. 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, see [15] for more background.

Theorem 1.2: Given $s \in (-1/d, 0]$ and $r \in (-sd, 1)$, there exists a constant $c = c(s, r, d, n)$ such that for all independent random vectors $X_1, \ldots, X_n$ in $\mathbb{R}^d$ with s-concave densities,

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

In particular, one can take

$$c = r^{\frac{1}{1-r}} \left(1 + \frac{1}{n|r'|}\right)^{1+n|r'|} B_1(s),$$

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

$$B_1(s) = \left(\prod_{k=1}^{d} \frac{(1+ks)^{|r'|(n-1)}(1+\frac{ks}{r})^{1+|r'|}}{(1+ks(1+\frac{1}{n|r'|}))^{1+n|r'|}} \right)^{\frac{2}{d}}.$$

Theorem 1.3: Let $s \in (-1/d, 0]$. There exist $0 < r_0 < 1$ and $\alpha = \alpha(s, r, d, n)$ such that for $r_0 \leq r < 1$ 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

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

and

$$\alpha = \left(1 + \frac{\log r + (r+1)\log\frac{r+1}{2r} + B_2(s)}{(1-r)\log 2}\right)^{-1},$$

where

$$B_2(s) = \frac{2}{d} \sum_{k=1}^{d} \left(\log\left(1 + \frac{ks}{r}\right) + r\log(1+ks) - (r+1)\log\left(1 + \frac{ks(r+1)}{2r}\right)\right).$$

Set $s = 0$ in Theorem 1.2 and Theorem 1.3. Then one can recover the log-concave case in [18]. The reader is directed to the full paper expanding on this work [13].

II. Proof Ingredients

It was first observed by Lieb [14] that the classical EPI can be proved via establishing an equivalent linearization form. Our proofs of Theorem 1.2 and Theorem 1.3 follow this approach. The following linearization of (3) and (4) is due to Rioul [21]. The case $c = 1$ has been used in [11].

Theorem 2.1 ([21]): Let $X_1, \ldots, X_n$ be independent random vectors in $\mathbb{R}^d$. The following statements are equivalent.

- There exist a constant $c > 0$ and an exponent $\alpha > 0$ such that
$$N_r^\alpha\left(\sum_{i=1}^{n} X_i\right) \geq c \sum_{i=1}^{n} N_r^\alpha(X_i).$$
- For any $\lambda_1, \ldots, \lambda_n \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$, one has
$$h_r\left(\sum_{i=1}^{n} \sqrt{\lambda_i} X_i\right) - \sum_{i=1}^{n} \lambda_i h_r(X_i) \geq \frac{d}{2}\left(\frac{\log c}{\alpha} + \left(\frac{1}{\alpha} - 1\right) H(\lambda)\right), \quad (7)$$
where $H(\lambda) \triangleq H(\lambda_1, \ldots, \lambda_n)$ is the discrete entropy defined as
$$H(\lambda) = -\sum_{i=1}^{n} \lambda_i \log \lambda_i.$$

One of the ingredients used to establish (7) is Young's sharp convolution inequality [1], [7]. Its information-theoretic formulation was given in [9], which we recall below. We denote by r' the Hölder conjugate of r, i.e.,

$$\frac{1}{r} + \frac{1}{r'} = 1.$$

Theorem 2.2 ([1], [7], [9]): Let $r > 0$. Let $\lambda_1, \ldots, \lambda_n \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$, and let $r_1, \ldots, r_n$ be positive reals such that $\lambda_i = r'/r'_i$. For independent random vectors $X_1, \ldots, X_n$ in $\mathbb{R}^d$, we have

$$h_r\left(\sum_{i=1}^{n} \sqrt{\lambda_i} X_i\right) - \sum_{i=1}^{n} \lambda_i h_{r_i}(X_i) \geq \frac{d}{2} r' \left(\frac{\log r}{r} - \sum_{i=1}^{n} \frac{\log r_i}{r_i}\right). \quad (8)$$

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 (7) from (8) 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, Li, and Madiman [10] serves this purpose.

Theorem 2.3 ([10]): Let $s \in \mathbb{R}$. Let $f\colon \mathbb{R}^d \to [0, +\infty)$ be an integrable s-concave function. Then, the function

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

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

$$C(r) = (r+s) \cdots (r+sd).$$

We deduce the following Rényi comparison for s-concave random variables.

Corollary 2.1: Let X be a random vector in $\mathbb{R}^d$ with 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 $q = (1-\lambda) \cdot r + \lambda \cdot 1$. By Theorem 2.3, we have

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

Rewrite the above inequality in terms of entropy power

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

The desired result follows from taking the logarithm of both sides. ■

Note that the condition $-sd < r < 1$ in Corollary 2.1 implies $s > -1/d$.

By combining Theorem 2.2 and Corollary 2.1, we can establish the following Rényi entropy power inequality valid for a single Rényi parameter $r \in (0,1)$ in the class of s-concave random variables.

Theorem 2.4: Let $s \in (-1/d, 0]$ and $r \in (-sd, 1)$. Let $X_1, \ldots, X_n$ be independent random vectors in $\mathbb{R}^d$ with s-concave densities. Then, for all $\lambda = (\lambda_1, \ldots, \lambda_n) \in [0,1]^n$ such that $\sum_{i=1}^n \lambda_i = 1$, one has

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

where

$$A(\lambda) = r'\left(1-\frac{1}{r'}\right)\log\left(1-\frac{1}{r'}\right) - r'\sum_{i=1}^n\left(1-\frac{\lambda_i}{r'}\right)\log\left(1-\frac{\lambda_i}{r'}\right), \tag{9}$$

and

$$g_k(\lambda) = (1-n)r'\log(1+ks) + (1-r')\log\left(1+\frac{ks}{r}\right) + r'\sum_{i=1}^n\left(1-\frac{\lambda_i}{r'}\right)\log\left(1+ks\left(1-\frac{\lambda_i}{r'}\right)\right). \tag{10}$$

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 2.2 with Corollary 2.1, we have

$$h_r\left(\sum_{i=1}^n \sqrt{\lambda_i}X_i\right) - \sum_{i=1}^n \lambda_i h_r(X_i) \geq \frac{dr'}{2}\left(\frac{\log r}{r} - \sum_{i=1}^n \frac{\log r_i}{r_i}\right) + \sum_{i=1}^n \lambda_i \log\frac{C(r)^{\frac{1}{1-r}}C(1)^{\frac{r_i-r}{(1-r_i)(1-r)}}}{C(r_i)^{\frac{1}{1-r_i}}}. \tag{11}$$

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

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

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

$$\begin{aligned}
&\sum_{i=1}^n \lambda_i\left(\frac{\log D(r)}{1-r} + \left(\frac{1}{1-r_i} - \frac{1}{1-r}\right)\log D(1) - \frac{\log D(r_i)}{1-r_i}\right) \\
&= (1-r')\log D(r) + (1-n)r'\log D(1) + \sum_{k=1}^d\sum_{i=1}^n (r'-\lambda_i)\log\left(1+\frac{ks}{r_i}\right) \\
&= \sum_{k=1}^d \left[(1-r')\log\left(1+\frac{ks}{r}\right) + (1-n)r'\log(1+ks) + \sum_{i=1}^n (r'-\lambda_i)\log\left(1+\frac{ks}{r_i}\right)\right] \\
&= \sum_{k=1}^d g_k(\lambda),
\end{aligned}$$

the last identity follows from $1/r_i = 1 - \lambda_i/r'$. Hence, the RHS of (11) can be written as

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

■

III. Proofs

Having Theorem 2.1 and Theorem 2.4 at hand, we are ready to prove the main results.

A. Proof of Theorem 1.2

Combine Theorem 2.1 with Theorem 2.4. Then it suffices to determine c such that for all $\lambda = (\lambda_1, \ldots, \lambda_n) \in [0,1]^n$ satisfying $\sum_{i=1}^n \lambda_i = 1$,

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

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 $\lambda = (\lambda_1, \ldots, \lambda_n) \in [0,1]^n$ such that $\sum_{i=1}^n \lambda_i = 1$. For fixed r, both $A(\lambda)$ and $g_k(\lambda)$ are the 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 value of c in Theorem 1.2.

B. Proof of Theorem 1.3

First, we state a lemma in [18], which will be used in the proof of Theorem 1.3.

Lemma 3.1 ([18]): 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 1.3: Using Theorem 2.1 and theorem 2.4 with $n = 2$, it suffices to find α such that for all $\lambda \in [0, 1]$,

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

where,

$$A(\lambda) = r'\left(1-\frac{1}{r'}\right)\log\left(1-\frac{1}{r'}\right) - r'\left(1-\frac{\lambda}{r'}\right)\log\left(1-\frac{\lambda}{r'}\right) - r'\left(1-\frac{1-\lambda}{r'}\right)\log\left(1-\frac{1-\lambda}{r'}\right),$$

and

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

We can set

$$\alpha = \left(1 - \sup_{0\leq\lambda\leq1}\left(-\frac{A(\lambda)}{H(\lambda)} - \frac{2}{d}\sum_{k=1}^{d}\frac{g_k(\lambda)}{H(\lambda)}\right)\right)^{-1}. \tag{12}$$

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

$$-\frac{A(\lambda)}{H(\lambda)} - \frac{2}{d}\sum_{k=1}^{d}\frac{g_k(\lambda)}{H(\lambda)} \tag{13}$$

is increasing on $[0, 1/2]$. It has been shown in [11] that $-A(\lambda)/H(\lambda)$ is increasing on $[0, 1/2]$. We will show that every $-g_k(\lambda)/H(\lambda)$ is also increasing on $[0, 1/2]$, by applying Lemma 3.1. Note that $-g_k(\lambda), H(\lambda) \geq 0$. Also, one can check that $g_k(0) = g_k(1) = 0$ and $g'_k(1/2) = 0$. Elementary calculations yield

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

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

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

Hence, we have

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

where

$$W(x) = xy\left(\frac{1}{1+ks(1+x)} + \frac{1}{1+ks(1+y)} + \frac{1}{(1+ks(1+x))^2} + \frac{1}{(1+ks(1+y))^2}\right) \tag{14}$$

with $y = \frac{1}{|r'|} - x$. Since $s, r' < 0$, it suffices to show that $W(x)$ is increasing over $[0, \frac{1}{2|r'|}]$. We rewrite W in the following way

$$W(x) = W_1(x) + W_2(x), \tag{15}$$

where

$$W_1(x) = xy\left(\frac{1}{1+ks(1+x)} + \frac{1}{1+ks(1+y)}\right), \tag{16}$$

and

$$W_2(x) = xy\left(\frac{1}{(1+ks(1+x))^2} + \frac{1}{(1+ks(1+y))^2}\right). \tag{17}$$

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 $y = \frac{1}{|r'|} - x$, it is easy to see 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

$$a \triangleq a(x) = 1 + ks(1 + x) \tag{18}$$

$$b \triangleq b(x) = 1 + ks(1 + y) = 1 + ks\left(\frac{1}{|r'|} - x + 1\right). \tag{19}$$

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

$$\begin{aligned} W'_1(x) &= \left(\frac{1}{a} + \frac{1}{b}\right)\left(\frac{1}{|r'|} - 2x - ksxy\left(\frac{1}{a} - \frac{1}{b}\right)\right) \\ &= \left(\frac{1}{a} + \frac{1}{b}\right)\left(\frac{1}{|r'|} - 2x\right)\left(1 - (ks)^2\frac{xy}{ab}\right), \end{aligned}$$

where 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 (18) and (19), 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 (17), it is easy to check that

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

where

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

Since

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

it suffices to show that $T(x) \geq 0$ for $x \in [0, 1/(2|r'|)]$. Using the identities

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

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

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

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 1.3 follows from setting $\lambda = 1/2$ in (12). ■

Remark 3.1: The key insight of the optimization argument in the proof is the monotonicity of $-A(\lambda)/H(\lambda)$ and $-g_k(\lambda)/H(\lambda)$ over $\lambda \in [0, 1/2]$. The monotonicity of $-A(\lambda)/H(\lambda)$ is independent of r. Numerical examples show that $-g_k(\lambda)/H(\lambda)$, even the optimization quantity in (13), is not monotone when r is small. This is one of the reasons for the restriction $r > r_0$.

Remark 3.2: Note that the condition $r > -sd$ of Theorem 2.3 can be rewritten as

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

We do not know whether Theorem 1.3 holds when

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

Acknowledgement

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

References

[1] W. Beckner. Inequalities in Fourier analysis. *Ann. of Math. (2)*, 102(1):159–182, 1975.

[2] 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.

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

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

[5] C. Borell. Convex measures on locally convex spaces. *Ark. Mat.*, 12:239–252, 1974.

[6] C. Borell. Convex set functions in *d*-space. *Period. Math. Hungar.*, 6(2):111–136, 1975.

[7] H. J. Brascamp and E. H. Lieb. Best constants in Young's inequality, its converse, and its generalization to more than three functions. *Advances in Math.*, 20(2):151–173, 1976.

[8] M. H. M. Costa. A new entropy power inequality. *IEEE Trans. Inform. Theory*, 31(6):751–760, 1985.

[9] A. Dembo, T. M. Cover, and J. A. Thomas. Information-theoretic inequalities. *IEEE Trans. Inform. Theory*, 37(6):1501–1518, 1991.

[10] M. Fradelizi, J. Li, and M. Madiman. Concentration of information content for convex measures. [Online]. Available: https://arxiv.org/abs/1512.01490.

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

[12] J. Li, A. Marsiglietti, and J. Melbourne. Entropic Central Limit Theorem for Rényi Entropy. *IEEE International Symposium on Information Theory*, Paris, France, 2019.

[13] J. Li, A. Marsiglietti, and J. Melbourne. Further investigations of Rényi entropy power inequalities and an entropic characterization of *s*-concave densities. [Online]. Available: https://arxiv.org/abs/1901.10616.

[14] E. H. Lieb. Proof of an entropy conjecture of Wehrl. *Comm. Math. Phys.*, 62(1):35–41, 1978.

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

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

[17] M. Madiman, J. Melbourne, and P. Xu. Infinity-Rényi entropy power inequalities *IEEE International Symposium on Information Theory* 2017.

[18] 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.

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

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

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

[22] G. Savaré and G. Toscani. The concavity of Rényi entropy power. *IEEE Trans. Inform. Theory*, 60(5):2687–2693, May 2014.

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

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