ENTROPIES OF SUMS OF INDEPENDENT GAMMA RANDOM VARIABLES

GIORGOS CHASAPIS, SALIL SINGH, AND TOMASZ TKOCZ

Abstract. We establish several Schur-convexity type results under fixed variance for weighted sums of independent gamma random variables and obtain nonasymptotic bounds on their Rényi entropies. In particular, this pertains to the recent results by Bartczak-Nayar-Zwara as well as Bobkov-Naumov-Ulyanov, offering simple proofs of the former and extending the latter.

2010 Mathematics Subject Classification. Primary 60E15; Secondary 94A17.

Key words. Entropy, max-entropy, gamma distribution, weighted sums, Schur-convexity

1. Introduction

Suppose $X_1, X_2, \ldots, X_n$ are independent, identically distributed (i.i.d.) square-integrable random variables, say with variance 1 and $a = (a_1, a_2, \ldots, a_n)$ is a unit vector in $\mathbb{R}^n$, $\sum_{k=1}^n a_k^2 = 1$, so that the variance of the sum $X_a = \sum_{k=1}^n a_k X_k$ does not depend on a and also equals 1. For which vectors a, is the distribution of X_a as close to the Gaussian distribution as possible? A natural way to quantify this vague question is to measure the distance to Gaussianity via relative entropy and ask about $\inf_a D(X_a||G)$. Here $D(X||G) = h(G) - h(X)$ is the relative entropy of X with respect to a Gaussian random variable G of the same variance as X, where

$$h(X) = -\int_{\mathbb{R}} f \log f$$

is the Shannon entropy of a random variable X with density f.

This question was raised in [6] and addressed for (symmetric) *Gaussian mixtures*, where the extremising sequence turns out to be simply $a = (\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})$. In a recent paper [1], Bartczak, Nayar and Zwara considered the case of gamma distribution and established that the same vector is extremal among all nonnegative vectors, that is whose all components are nonnegative. We refer to their paper for a comprehensive account of relevant reference and related problems. Their approach rests on the so-called method of interlacing densities (see also [7]). For the

Date: January 21, 2022.

TT's research supported in part by NSF grant DMS-1955175.

gamma distribution, this entails a rather technical and involved analysis for Bessel functions.

Our first goal in this paper is to offer an alternative approach. It turns out that for the gamma distribution, simple arguments involving moment generating functions allow to establish certain Schur-convexity type results. Those in particular give the main result of [1], as well as partially address Question 6 from [1] about moments.

Our second goal in this paper is to extend a recent result of [4], where Bobkov, Naumov and Ulyanov find a nonasymptotic expression for the maximum of the density of $X_a = \sum a_k X_k$ with X_k having $\Gamma(1/2)$ distribution, equivalently for the ∞-Rényi entropy of X_a in terms of a. We extend this to $\Gamma(\gamma)$ distribution with $\gamma \geq 1/2$. Such bounds have applications to Lévy's concentration function, thus to anti-concentration inequalities (see, e.g. [3] as well as e.g. the survey [11] for an exposition on anti-concentration).

Another piece of motivation to study such extensions is the fact that weighted sums of independent $\Gamma(1/2)$ random variables emerge naturally from Gaussian quadratic forms, which was a starting point for both [1] and [4].

In the next section, we recall the definition of Rényi entropy and formulate our results. The remaining part of this note will be devoted to their proofs.

Acknowledgements. We are indebted to Han Nguyen for many fruitful and illuminating discussions.

2. Results

Let $0 \leq \alpha \leq \infty$. For a random variable X with density f, we define its Rényi entropy of order α as

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

understood as limits in the cases $\alpha \in \{0, 1, \infty\}$, namely $h_0(X) = \log|\mathrm{supp}(f)|$, $h_1(X) = h(X)$ (the Shannon entropy), $h_\infty(X) = -\log\|f\|_\infty$. For notational convenience, we also introduce the functional

$$M(X) = \|f\|_\infty.$$

Throughout, we let $\gamma > 0$ and let $X_1, X_2, \ldots$ be i.i.d. random variables with $\Gamma(\gamma)$ distribution, that is with density $\Gamma(\gamma)^{-1} x^{\gamma-1} e^{-x} \mathbf{1}_{(0,\infty)}(x)$ on $\mathbb{R}$.

2.1. **Schur-convexity type results and entropies.** Our first main result gives the Schur-concavity of *centred* weighted sums averaged against arbitrary completely monotone functions. For concise exposition on majorisation and Schur-convexity, we refer for instance to Chapter II of [2]. We recall that a function $\Phi\colon (0, +\infty) \to (0, +\infty)$ is completely monotone if it is a *mixture* of exponential functions, that is $\Phi(x) = \int_0^\infty e^{-tx} \mathrm{d}\mu(t)$ for some nonnegative Borel measure μ, equivalently (by Bernstein's theorem) $(-1)^m \Phi^{(m)}(x) \geq 0$ for every $m = 0, 1, 2, \ldots$, see, e.g. [8].

Theorem 1. *For a completely monotone function $\Phi\colon (0, +\infty) \to (0, +\infty)$ and $c > 0$, the function*

$$(a_1, \ldots, a_n) \mapsto \mathbb{E}\Phi\left(c + \sum_{j=1}^{n} \sqrt{a_j}(X_j - \gamma)\right) \tag{1}$$

is Schur-concave on the simplex $\{a \in \mathbb{R}_+^n,\ \sum a_j < \frac{c^2}{\gamma^2 n}\}$.

We emphasise that the centering of the X_j by its mean $\mathbb{E}X_j = \gamma$ is crucial for this result to hold. Without the centering, the resulting function is Schur-convex, as will follow from our proof.

Theorem 2. *For a completely monotone function $\Phi\colon (0, +\infty) \to (0, +\infty)$, the function*

$$(a_1, \ldots, a_n) \mapsto \mathbb{E}\Phi\left(\sum_{j=1}^{n} \sqrt{a_j} X_j\right) \tag{2}$$

is Schur-convex on $\mathbb{R}_+^n$.

The main result of [1] follows as a corollary to Theorem 1.

Corollary 3 (Bartczak-Nayar-Zwara, [1])**.** *Provided that $\gamma n \geq 1$, we have for the Shannon entropy,*

$$h\left(\sum_{j=1}^{n} \sqrt{a_j} X_j\right) \leq h\left(\sum_{j=1}^{n} \frac{1}{\sqrt{n}} X_j\right), \tag{3}$$

whenever $\sum a_j = 1$.

For general α-Rényi entropies, we can deduce the same, but using Theorem 2 and imposing additional restrictions on the parameters α, γ and n. This can be compared with results for Gaussian mixtures (Theorem 8 in [6]) as well as sums of uniform random variables (Theorem 2 in [5]).

Corollary 4. *Let $\alpha > 1$, $n\gamma < 1$ and $\sum_{j=1}^{n} a_j = 1$. We have,*

$$h_\alpha\left(\sum_{j=1}^{n} \sqrt{a_j} X_j\right) \leq h_\alpha\left(\sum_{j=1}^{n} \frac{1}{\sqrt{n}} X_j\right). \tag{4}$$

We also have Schur-convexity for power functions with integral exponents. This relates to Question 6 from [1], except that here we are only able to handle *centred* moments of even order.

Theorem 5. *For every positive integer k, the function*

$$(a_1, \dots, a_n) \mapsto \mathbb{E}\left(\sum_{j=1}^{n} \sqrt{a_j}(X_j - \gamma)\right)^k \tag{5}$$

is Schur-convex on $\mathbb{R}_+^n$.

2.2. **Maximum density.** Let $a_1 \geq \cdots \geq a_n > 0$, $\sum_{j=1}^n a_j = 1$. The main result of Bobkov-Naumov-Ulyanov from [4] asserts that when $\gamma = 1/2$ (i.e. when X_j has the same distribution as $\frac{1}{2}Z^2$), we have

$$\frac{1}{2e^2\sqrt{2\pi}}(1-a_1)^{-1/4} \leq M\left(\sum_{j=1}^{n} \sqrt{a_j} X_j\right) \leq \frac{4}{\sqrt{\pi}}(1-a_1)^{-1/4}, \tag{6}$$

Using their approach, we extend this to $\gamma \geq 1/2$. For $\gamma \geq 1$, M is of the constant order, $\gamma^{-1/2}$ up to universal constants. For $\frac{1}{2} \leq \gamma < 1$, only the exponent in (6) has to be modified (and of course the universal constants).

Theorem 6. *For $\gamma \geq 1$, we have*

$$\frac{1}{\sqrt{12}}\gamma^{-1/2} \leq M\left(\sum_{j=1}^{n} \sqrt{a_j} X_j\right) \leq \frac{e}{\sqrt{\pi}}\gamma^{-1/2}. \tag{7}$$

For $\frac{1}{2} \leq \gamma < 1$, there are constants c_γ and C_γ for which we have

$$c_\gamma(1-a_1)^{\frac{\gamma-1}{2}} \leq M\left(\sum_{j=1}^{n} \sqrt{a_j} X_j\right) \leq C_\gamma(1-a_1)^{\frac{\gamma-1}{2}}. \tag{8}$$

The lower bound in fact holds for every $0 < \gamma < 1$ and we can take $c_\gamma = 0.003\gamma$.

We will prove slightly more general results, allowing to justify the following remark.

Remark 7. For $\gamma < 1$, $M = +\infty$ regardless of a as long as $n \leq \lfloor 1/\gamma \rfloor$ (e.g. see Lemma 13 below). For $\gamma < \frac{1}{2}$, we only know the matching lower and upper bounds when $n = \lfloor 1/\gamma \rfloor + 1$ (see Remark 15 in the next section). The case of arbitrary n has been elusive and we find it an interesting question.

3. Proofs

We note for future use the formula for the moment generating function of a $\Gamma(\gamma)$ random variable X: for $t < 1$,

$$\mathbb{E}e^{tX} = (1-t)^{-\gamma}. \tag{9}$$

3.1. **Proof of Theorems 1 and 2.** We begin with a lemma.

Lemma 8. *Let $\gamma > 0$. The function $F(x_1, \ldots, x_n) = \prod_{j=1}^n e^{\gamma\sqrt{x_j}}(1 + \sqrt{x_j})^{-\gamma}$ is Schur-concave on $\mathbb{R}_+^n$, whereas the function $G(x_1, \ldots, x_n) = \prod_{j=1}^n (1 + \sqrt{x_j})^{-\gamma}$ is Schur-convex on $\mathbb{R}_+^n$.*

Proof. For function F, we have,

$$\begin{aligned}\frac{1}{F(x)}\frac{\partial F}{\partial x_k} = \frac{\partial}{\partial x_k}\log F &= \gamma\frac{\partial}{\partial x_k}\left(\sqrt{x_k} - \log(1+\sqrt{x_k})\right) \\ &= \frac{\gamma}{2\sqrt{x_k}}\left(1 - \frac{1}{1+\sqrt{x_k}}\right) \\ &= \frac{\gamma}{2}\frac{1}{1+\sqrt{x_k}}.\end{aligned}$$

Thus, if $x_k > x_l$, then

$$\frac{\partial F}{\partial x_k} - \frac{\partial F}{\partial x_l} = \frac{\gamma}{2}F(x)\left(\frac{1}{1+\sqrt{x_k}} - \frac{1}{1+\sqrt{x_l}}\right) < 0.$$

The Schur-Ostrowski criterion finishes the proof for F. For function G, the argument proceeds identically. □

Proof of Theorem 1. Since Φ is completely monotone, there is a (nonnegative) Borel measure on μ such that

$$\Phi(x) = \int_0^\infty e^{-tx}\mathrm{d}\mu(t).$$

Thus, thanks to independence and (9),

$$\mathbb{E}\Phi\left(c + \sum_{j=1}^n \sqrt{a_j}(X_j - \gamma)\right) = \int_0^\infty \left(\prod_{j=1}^n e^{\gamma t\sqrt{a_j}}(1 + t\sqrt{a_j})^{-\gamma}\right) e^{-ct}\mathrm{d}\mu(t).$$

Lemma 8 finishes the proof. □

Remark 9. We emphasise that the factor $e^{\gamma\sqrt{a_j}}$ appears as a result of centreing the X_j. This factor is crucial for function F from Lemma 8 to be Schur-concave, as without it, as we have seen, it is Schur-convex. Theorem 2 follows analogously.

3.2. **Proof of Corollary 3.** First note that applying Theorem 1 to $\Phi(x) = x^{-q}$ with $q \to 0+$ and using that $\frac{x^{-q}-1}{q} \downarrow -\log x$, as $q \downarrow 0+$ for positive x, we conclude that Theorem 1 also holds with $\Phi(x) = -\log x$. To prove (3), fix $c > \gamma\sqrt{n}$ and positive a_j with $\sum_{j=1}^n a_j = 1$. Recall that for an arbitrary probability density function g, we have

$$\begin{aligned}h\left(\sum_{j=1}^n \sqrt{a_j}X_j\right) &= h\left(c + \sum_{j=1}^n \sqrt{a_j}(X_j - \gamma)\right) \\ &\leq \mathbb{E}\left[-\log g\left(c + \sum_{j=1}^n \sqrt{a_j}(X_j - \gamma)\right)\right].\end{aligned}$$

Letting g be the density of $\sum_{j=1}^n \frac{1}{\sqrt{n}} X_j$, that is

$$g(x) = \frac{n^{(\gamma n-1)/2}}{\Gamma(\gamma n)} x_+^{\gamma n-1} e^{-x\sqrt{n}},$$

we thus obtain from the first part that (note that we need $\gamma n - 1 \geq 0$)

$$h\left(\sum_{j=1}^n \sqrt{a_j} X_j\right) \leq \mathbb{E}\left[-\log g\left(c + \sum_{j=1}^n \frac{1}{\sqrt{n}}(X_j - \gamma)\right)\right]$$

With $c \to \gamma\sqrt{n}+$, the right hand side becomes $h\left(\sum_{j=1}^n \frac{1}{\sqrt{n}} X_j\right)$. □

3.3. **Proof of Corollary 4.** Suppose $\sum a_j = 1$, let f be the density of $\sum \sqrt{a_j} X_j$ and g be the density of $\sum X_j/\sqrt{n}$. Our goal is to show that $\int f^\alpha \geq \int g^\alpha$. By Hölder's inequality,

$$\left(\int f^\alpha\right)^{\frac{1}{\alpha}} \left(\int g^\alpha\right)^{\frac{\alpha-1}{\alpha}} \geq \int f g^{\alpha-1}.$$

Note that the right hand side reads $\mathbb{E}\Phi(\sum \sqrt{a_j} X_j)$ with

$$\Phi(x) = g(x)^{\alpha-1} = \left(\Gamma(n\gamma)^{-1}(x\sqrt{n})^{n\gamma-1} e^{-x\sqrt{n}}\right)^{\alpha-1}$$

which is completely monotone as a product of two completely monotone functions (hence the assumptions, to have $(n\gamma-1)(\alpha-1) < 0$ and $\alpha - 1 > 0$). It remains to apply Theorem 2 to the sequence (a_j) which always majorises the constant sequence $(\frac{1}{n})$. □

3.4. **Proof of Theorem 5.** First we prove a lemma about centred integral moments of a single summand.

Lemma 10. *For every positive integer k, $\mathbb{E}(X_1 - \gamma)^k \geq 0$.*

Proof. Rephrasing the lemma, it suffices to prove that the power-series expansion of the moment generating function $\mathbb{E}e^{t(X_1-\gamma)}$ has nonnegative coefficients. Invoking (9) and using that $-\log(1-t) = \sum_{k=1}^\infty \frac{t^k}{k}$, we obtain

$$\mathbb{E}e^{t(X_1-\gamma)} = \exp\left\{\gamma\left(-t - \log(1-t)\right)\right\} = \exp\left\{\gamma \sum_{k=2}^\infty \frac{t^k}{k}\right\}.$$

Since the power series expansion of exp has positive coefficients, the proof is complete. □

Proof of Theorem 5. Let $S = \sum_{j=1}^{n} \sqrt{a_j}(X_j - \gamma)$. Consider for sufficiently small positive t,

$$\begin{aligned}\mathbb{E}e^{tS} &= \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbb{E}S^k \\ &= \prod_{j=1}^{n} \exp\left\{-\gamma\Big(t\sqrt{a_j} + \log(1 - t\sqrt{a_j})\Big)\right\}.\end{aligned}$$

Call the right hand side F. Fix two indices $i \neq j$ and observe that $\left(\frac{\partial}{\partial a_j} - \frac{\partial}{\partial a_i}\right)\mathbb{E}S^k$ is the Taylor coefficient of $\left(\frac{\partial}{\partial a_j} - \frac{\partial}{\partial a_i}\right) F$ at t^k. On the other hand,

$$\frac{\partial F}{\partial a_j} = F \cdot \left(-\gamma\left(\frac{t}{2\sqrt{a_j}} - \frac{t}{2\sqrt{a_j}(1 - t\sqrt{a_j})}\right)\right) = F \cdot \frac{\gamma t^2}{2(1 - t\sqrt{a_j})}.$$

Thus,

$$\left(\frac{\partial}{\partial a_j} - \frac{\partial}{\partial a_i}\right) F = F \cdot \frac{\gamma t^3}{2(1 - t\sqrt{a_j})(1 - t\sqrt{a_i})}(\sqrt{a_j} - \sqrt{a_i}).$$

For $a_j > a_i$, the power-series expansion of the right hand side has nonnegative coefficients (F has, by Lemma 10, and plainly so does $(1 - t\sqrt{a_j})^{-1}$). Combining this with the Schur-Ostrowski criterion finishes the argument. □

3.5. **Proof of Theorem 6.** We assume throughout that $a_1 \geq a_2 \geq \ldots$. For the proofs, we recall two lemmas: the first one is classical and goes back to Moriguti and the second one is a straightforward extension of Lemma 3 from [4].

Lemma 11 (Moriguti, [10])**.** *For every random variable X, $M(X) \geq \frac{1}{\sqrt{12}} \frac{1}{\sqrt{\mathrm{Var}(X)}}$.*

Lemma 12. *If $a_1 \leq \frac{1}{m}$ for a positive integer m, then the characteristic function ϕ of $\sqrt{a_1}X_1 + \cdots + \sqrt{a_n}X_n$ satisfies*

$$|\phi(t)| \leq (1 + t^2/m)^{-m\gamma/2}, \qquad t \in \mathbb{R}. \tag{10}$$

Moreover, if $m\gamma > 1$,

$$M\left(\sqrt{a_1}X_1 + \cdots + \sqrt{a_n}X_n\right) \leq \frac{\sqrt{m}\Gamma\left(\frac{m\gamma - 1}{2}\right)}{2\sqrt{\pi}\Gamma\left(\frac{m\gamma}{2}\right)}. \tag{11}$$

Proof. The characteristic function of X_1 is

$$\mathbb{E}e^{itX_1} = (1 - it)^{-\gamma}, \qquad t \in \mathbb{R},$$

with choosing, say the principal branch. Thus,

$$\phi(t) = \prod_{j=1}^{n}(1 - i\sqrt{a_j}t)^{-\gamma}$$

and

$$\log|\phi(t)| = -\frac{\gamma}{2}\sum_{j=1}^{n}\log(1 + a_j t^2).$$

To finish the proof of (10), we find the maximum of the convex function

$$(a_1, \dots, a_n) \mapsto -\sum_{j=1}^{n} \log(1 + a_j t^2)$$

over the domain $D = \{(a_1, \dots, a_n),\ a_1, \dots, a_n \geq 0,\ a_1 + \dots + a_n = 1\} \cap [0, 1/m]^n$. We can either follow [4] verbatim and examine its extreme points, or, alternatively, it is clear that an arbitrary vector in D is majorised by the vector $(\frac{1}{m}, \dots, \frac{1}{m}, 0, \dots, 0)$ (with $\frac{1}{m}$ repeated m-times and 0 repeated $n-m$ times), so the lemma follows from the Schur-convexity of this function.

To see (11), we apply the Fourier inversion formula and (10),

$$\begin{aligned} M\left(\sqrt{a_1}X_1 + \dots + \sqrt{a_n}X_n\right) \leq \frac{1}{2\pi}\int_{\mathbb{R}} |\phi(t)|\mathrm{d}t &\leq \frac{1}{2\pi}\int_{\mathbb{R}} (1 + t^2/m)^{-m\gamma/2}\mathrm{d}t \\ &= \frac{\sqrt{m}\Gamma\left(\frac{m\gamma - 1}{2}\right)}{2\sqrt{\pi}\Gamma\left(\frac{m\gamma}{2}\right)}. \end{aligned}$$

□

We will also need a simple point-wise bound on the density of the sum $\sqrt{a_1}X_1 + \dots + \sqrt{a_n}X_n$.

Lemma 13. *The density p of $\sqrt{a_1}X_1 + \dots + \sqrt{a_n}X_n$ satisfies*

$$\frac{1}{\Gamma(n\gamma)}(a_1 \dots a_n)^{-\gamma/2} x^{n\gamma - 1} e^{-x/\sqrt{a_n}} \leq p(x) \leq \frac{1}{\Gamma(n\gamma)}(a_1 \dots a_n)^{-\gamma/2} x^{n\gamma - 1}.$$

Proof. Fix $x > 0$. By independence, convolving the densities of $\sqrt{a_j}X_j$ yields

(12)

$$\begin{aligned} p(x) =& \Gamma(\gamma)^{-n}(a_1 \dots a_n)^{-\gamma/2} \\ & \int_{\substack{t_1, \dots, t_{n-1} > 0, \\ t_1 + \dots + t_{n-1} < x}} \Bigg[(t_1 \dots t_{n-1})^{\gamma - 1}(x - t_1 - \dots - t_{n-1})^{\gamma - 1} \\ & \cdot \exp\left\{ -\frac{t_1}{\sqrt{a_1}} - \dots - \frac{t_{n-1}}{\sqrt{a_{n-1}}} - \frac{x - t_1 - \dots - t_{n-1}}{\sqrt{a_n}} \right\} \Bigg] \mathrm{d}t_1 \dots \mathrm{d}t_{n-1}. \end{aligned}$$

Changing each t_j to xt_j gives

(13)

$$\begin{aligned} p(x) =& \Gamma(\gamma)^{-n}(a_1 \dots a_n)^{-\gamma/2} x^{n\gamma - 1} \\ & \cdot \int_{\substack{t_1, \dots, t_{n-1} > 0, \\ t_1 + \dots + t_{n-1} < 1}} \Bigg[(t_1 \dots t_{n-1})^{\gamma - 1}(1 - t_1 - \dots - t_{n-1})^{\gamma - 1} \\ & \cdot \exp\left\{ -x\left(\frac{t_1}{\sqrt{a_1}} + \dots + \frac{t_{n-1}}{\sqrt{a_{n-1}}} + \frac{1 - t_1 - \dots - t_{n-1}}{\sqrt{a_n}} \right) \right\} \Bigg] \mathrm{d}t_1 \dots \mathrm{d}t_{n-1}. \end{aligned}$$

Note that for the t_j from the integral's domain,

$$0 \le \frac{t_1}{\sqrt{a_1}} + \cdots + \frac{t_{n-1}}{\sqrt{a_{n-1}}} + \frac{1 - t_1 - \cdots - t_{n-1}}{\sqrt{a_n}}$$
$$= \frac{1}{\sqrt{a_n}} + \sum_{j=1}^{n-1} t_j \left(\frac{1}{\sqrt{a_j}} - \frac{1}{\sqrt{a_n}} \right) \le \frac{1}{\sqrt{a_n}}$$

(recalling that $a_j \ge a_n$). The resulting estimates on $\exp\{\dots\}$ in the integrand give the desired bounds on p, where the factor $\frac{1}{\Gamma(n\gamma)}$ comes from

$$\Gamma(\gamma)^{-n} \int_{\substack{t_1,\dots,t_{n-1}>0,\\ t_1+\cdots+t_{n-1}<1}} (t_1 \dots t_{n-1})^{\gamma-1} (1 - t_1 - \cdots - t_{n-1})^{\gamma-1} \mathrm{d}t_1 \dots \mathrm{d}t_{n-1}$$
$$= \Gamma(\gamma)^{-n} B(\gamma,\gamma) B(2\gamma,\gamma) \dots B((n-1)\gamma,\gamma) = \frac{1}{\Gamma(n\gamma)}.$$

□

We also recall the following standard Stirling's approximation for the gamma function: for $x > 0$

$$\sqrt{2\pi} x^{x-1/2} e^{-x} \le \Gamma(x) \le \sqrt{2\pi} e^{1/(12x)} x^{x-1/2} e^{-x} \tag{14}$$

(see, e.g. [9]).

We move to the proof of Theorem 6. First we assume that $\gamma \ge 1$.

Proof of (7), *the upper bound.* Note that, with the aid of (14),

$$M(X_1) = \Gamma(\gamma)^{-1} \sup_{x>0} x^{\gamma-1} e^{-x} = \frac{(\gamma-1)^{\gamma-1} e^{-(\gamma-1)}}{\Gamma(\gamma)} \le \frac{(\gamma-1)^{\gamma-1} e^{-(\gamma-1)}}{\sqrt{2\pi} \gamma^{\gamma-1/2} e^{-\gamma}}$$
$$\le \frac{e}{\sqrt{2\pi}} \gamma^{-1/2}.$$

Since $M(X+Y) \le M(X)$ for every independent random variables X and Y, we have

$$M\left(\sum_{j=1}^{n} \sqrt{a_j} X_j \right) \le M(\sqrt{a_1} X_1) = a_1^{-1/2} M(X_1)$$

which gives the desired upper bound if $a_1 > \frac{1}{2}$, say. Otherwise, $a_1 \le \frac{1}{2}$, so (11) applied with $m = 2$ gives the upper bound of $\frac{\Gamma(\gamma-1/2)}{\sqrt{2\pi}\Gamma(\gamma)}$ which thanks to (14) is at most $\frac{e^{2/3}}{\sqrt{2\pi}} \gamma^{-1/2}$. □

Proof of (7), *the lower bound.* It immediately follows from Lemma 11 since we have, $\mathrm{Var}(\sum \sqrt{a_j} X_j) = \gamma$. □

Now we assume that $\gamma < 1$. The upper bound in (8) as well as Remark 7 follow from the following upper bound.

Theorem 14. *Fix a positive integer k and let $\frac{1}{k+1} \le \gamma < \frac{1}{k}$. Then*

$$M\left(\sum_{j=1}^{n} \sqrt{a_j} X_j\right) \le C_\gamma (a_1 \dots a_k)^{-\gamma/2} (1 - a_1 - \dots - a_k)^{\frac{k\gamma-1}{2}} \tag{15}$$

with the right hand side understood as $+\infty$ when $n \le k$. Constant C_γ depends only on γ.

Remark 15. In particular, when $k = 1$, this gives the upper bound in (8). Unfortunately, when $\gamma < \frac{1}{2}$, that is $k \ge 2$, bound (15) is not optimal: consider for instance the case when $a_1 = \dots = a_n = \frac{1}{n}$ with large n. However, note that when $n = k+1$, bound (15) is matched from below by Lemma 13 which gives that in this case

$$M\left(\sum_{j=1}^{k+1} \sqrt{a_j} X_j\right) \ge c_\gamma (a_1 \dots a_k)^{-\gamma/2} a_{k+1}^{\frac{k\gamma-1}{2}}$$

with $c_\gamma = \frac{((k+1)\gamma-1)^{(k+1)\gamma-1}}{e^{(k+1)\gamma-1}\Gamma((k+1)\gamma)}$, justifying Remark 7.

Proof of (15). In the course of the proof the value of C_γ may change from line to line. Note that when $n \le k$, by Lemma 13, the maximum of the density is $+\infty$ (because the exponent at x is negative). Thus, we can assume that $n \ge k+1$.

Case $n = k+1$. From (12), after changing the variables (scaling each t_i by $\sqrt{a_{k+1}}x$) we have,

(16)

$$\begin{aligned} p(\sqrt{a_{k+1}}x) &= \Gamma(\gamma)^{-k-1} (a_1 \dots a_k)^{-\gamma/2} a_{k+1}^{\frac{k\gamma-1}{2}} x^{(k+1)\gamma-1} \\ &\cdot \int_{\substack{t_1,\dots,t_k>0,\\ t_1+\dots+t_k<1}} \Bigg[(t_1 \dots t_k)^{\gamma-1} (1 - t_1 - \dots - t_k)^{\gamma-1} \\ &\quad \cdot \exp\left\{-x\left(\sqrt{\frac{a_{k+1}}{a_1}}t_1 + \dots + \sqrt{\frac{a_{k+1}}{a_k}}t_k + 1 - t_1 - \dots - t_k\right)\right\}\Bigg] \mathrm{d}t_1 \dots \mathrm{d}t_k. \end{aligned}$$

The crude bound $\sum \sqrt{\frac{a_{k+1}}{a_j}} t_j \ge 0$ yields

$$\begin{aligned} p(\sqrt{a_{k+1}}x) &\le \Gamma(\gamma)^{-k-1} (a_1 \dots a_k)^{-\gamma/2} a_{k+1}^{\frac{k\gamma-1}{2}} x^{(k+1)\gamma-1} \\ &\cdot \int_{\substack{t_1,\dots,t_k>0,\\ \sum t_j<1}} \left[(t_1 \dots t_k)^{\gamma-1} \left(1 - \sum t_j\right)^{\gamma-1} \cdot \exp\left\{-x\left(1 - \sum t_j\right)\right\}\right] \mathrm{d}t_1 \dots \mathrm{d}t_k. \end{aligned}$$

Using

$$x^{(k+1)\gamma-1} \exp\left\{-x\left(1 - \sum t_j\right)\right\} \le L_\gamma \left(1 - \sum t_j\right)^{1-(k+1)\gamma},$$

where $L_\gamma = \sup_{x>0} x^{(k+1)\gamma-1} e^{-x} = ((k+1)\gamma - 1)^{(k+1)\gamma-1} e^{-((k+1)\gamma-1)}$, we obtain the desired bound

$$\|p\|_\infty \le C_\gamma (a_1 \dots a_k)^{-\gamma/2} a_{k+1}^{\frac{k\gamma-1}{2}} \tag{17}$$

with

$$C_\gamma = \Gamma(\gamma)^{-k-1} L_\gamma \int_{\substack{t_1,\ldots,t_k>0,\\ \sum t_j<1}} (t_1\ldots t_k)^{\gamma-1}\left(1-\sum t_j\right)^{-k\gamma}\mathrm{d}t_1\ldots\mathrm{d}t_k$$

which is finite because $k\gamma < 1$.

Case $n \geq k+2$. If $a_1 \leq \frac{1}{k+2}$, then (11) applied to $m = k+2$ gives

$$M\left(\sum_{j=1}^n \sqrt{a_j}X_j\right) \leq C_\gamma \leq C_\gamma (a_1\ldots a_k)^{-\gamma/2}\left(1-a_1-\cdots-a_k\right)^{\frac{k\gamma-1}{2}},$$

since $a_1\ldots a_k \leq 1$, $1-a_1-\cdots-a_k \leq 1$, where C_γ only depends on γ. Now we assume that $a_1 > \frac{1}{k+2}$, write

$$\sum_{j=1}^n \sqrt{a_j}X_j = \sqrt{\alpha}\eta + \sqrt{1-\alpha}\xi$$

with

$$\alpha = \sum_{j=1}^k a_j$$

and

$$\eta = \sum_{j=1}^k \sqrt{\frac{a_j}{\alpha}}X_j, \qquad \xi = \sum_{j=k+1}^n \sqrt{\frac{a_j}{1-\alpha}}X_j.$$

We break the argument into two further cases depending on whether we can guarantee that ξ has a bounded density (using Lemma 12).

Case $\frac{a_{k+1}}{1-\alpha} \leq \frac{1}{k+2}$. Here necessarily the number of summands in ξ is at least $k+2$ (by comparing the largest coefficient to the average). Let g be the density of ξ. By (11) applied with $m = k+2$, we get $\|g\|_\infty \leq C_\gamma$. Moreover, if we let f be the density of η, we know by Lemma 13 that

$$f(x) \leq \frac{1}{\Gamma(k\gamma)}\alpha^{k\gamma/2}(a_1\ldots a_k)^{-\gamma/2}x^{k\gamma-1}, \qquad x>0.$$

Thus for the density p of $\sum_{j=1}^n \sqrt{a_j}X_j$, we obtain

$$\begin{aligned} p(x) &= \int_0^x \frac{1}{\sqrt{\alpha(1-\alpha)}} f\left(\frac{t}{\sqrt{\alpha}}\right) g\left(\frac{x-t}{\sqrt{1-\alpha}}\right)\mathrm{d}t \\ &\leq \frac{1}{\Gamma(k\gamma)\sqrt{1-\alpha}}(a_1\ldots a_k)^{-\gamma/2}\int_0^x t^{k\gamma-1} g\left(\frac{x-t}{\sqrt{1-\alpha}}\right)\mathrm{d}t. \end{aligned}$$

Changing x to $\sqrt{1-\alpha}x$ and t to $\sqrt{1-\alpha}t$, we get

$$p(\sqrt{1-\alpha}x) \leq \frac{1}{\Gamma(k\gamma)}(a_1\ldots a_k)^{-\gamma/2}(1-\alpha)^{\frac{k\gamma-1}{2}}\int_0^x t^{k\gamma-1} g\left(x-t\right)\mathrm{d}t.$$

It remains to observe that the resulting integral is bounded,

$$\int_0^x t^{k\gamma-1} g\left(x-t\right)\mathrm{d}t \leq \|g\|_\infty \int_{t<1} t^{k\gamma-1}\mathrm{d}t + \int_{t>1} g(x-t)\mathrm{d}t \leq \frac{\|g\|_\infty}{k\gamma} + 1.$$

Case $\frac{a_{k+1}}{1-\alpha} \geq \frac{1}{k+2}$. Plainly,

$$M\left(\sum_{j=1}^{n}\sqrt{a_j}X_j\right) \leq M\left(\sum_{j=1}^{k+1}\sqrt{a_j}X_j\right) = \frac{1}{\sqrt{A}}M\left(\sum_{j=1}^{k+1}\sqrt{\frac{a_j}{A}}X_j\right)$$

with $A = \sum_{j=1}^{k+1} a_j$. By the case $n = k+1$, i.e. (17),

$$M\left(\sum_{j=1}^{k+1}\sqrt{\frac{a_j}{A}}X_j\right) \leq C_\gamma(a_1 \dots a_k)^{-\gamma/2}A^{k\gamma/2}(a_{k+1}/A)^{\frac{k\gamma-1}{2}},$$

thus

$$\begin{aligned} M\left(\sum_{j=1}^{n}\sqrt{a_j}X_j\right) &\leq C_\gamma(a_1 \dots a_k)^{-\gamma/2}a_{k+1}^{\frac{k\gamma-1}{2}} \\ &\leq C_\gamma(k+2)^{-\frac{k\gamma-1}{2}}(a_1 \dots a_k)^{-\gamma/2}(1-\alpha)^{\frac{k\gamma-1}{2}}, \end{aligned}$$

as desired. This concludes the proof of (15). □

Proof of (8)*, the lower bound.* We assume that $0 < \gamma < 1$ and denote $Z = \sum \sqrt{a_j}X_j$. The argument from [4] can be repeated almost verbatim. We include it for completeness.

Case $a_1 \leq \frac{1}{2}$. Since $\mathrm{Var}(Z) = \gamma$, Lemma 11 yields $M(Z) \geq \frac{1}{2\sqrt{3\gamma}}$ so if $a_1 \leq 1/2$ then $M(Z) \geq c_\gamma(1-a_1)^{\frac{\gamma-1}{2}}$ with $c_\gamma = (2^{\frac{3-\gamma}{2}}\sqrt{3\gamma})^{-1}$.

Case $a_1 \geq 1/2$. Let $\xi = \sum_{j=2}^{n}\frac{\sqrt{a_j}}{\sqrt{1-a_1}}X_j$, so that $Z = \sqrt{a_1}X_1 + \sqrt{1-a_1}\xi$. Note that ξ is independent of $\sqrt{a_1}X_1$ so the density f_Z of Z is given by the convolution of the densities $f_{\sqrt{a_1}X_1}$ and $f_{\sqrt{1-a_1}\xi}$. We have,

$$f_Z(x) = \frac{1}{\Gamma(\gamma)\sqrt{a_1(1-a_1)}}\int_0^x (\frac{x-t}{\sqrt{a_1}})^{\gamma-1}\exp\left(-\frac{x-t}{\sqrt{a_1}}\right)f_\xi\left(\frac{t}{\sqrt{1-a_1}}\right)dt,$$

and, applying this for $x\sqrt{1-a_1}$,

$$f_Z(x\sqrt{1-a_1}) = \frac{(1-a_1)^{\frac{\gamma-1}{2}}}{\Gamma(\gamma)a_1^{\gamma/2}}\int_0^x (x-t)^{\gamma-1}\exp\left(-\frac{\sqrt{1-a_1}}{\sqrt{a_1}}(x-t)\right)f_\xi(t)\,dt.$$

We will use this identity for $x = \mathbb{E}\xi + 2$, lower bounding the expression in the right hand side by integrating on the interval $I = (\max(\mathbb{E}\xi - 2, 0), \mathbb{E}\xi + 2)$. Note that $x - t \leq 4$ for every $t \in I$. It follows that

$$M(Z) \geq \frac{(1-a_1)^{\frac{\gamma-1}{2}}}{\Gamma(\gamma)a_1^{\gamma/2}}4^{\gamma-1}\exp\left(-\frac{4\sqrt{1-a_1}}{\sqrt{a_1}}\right)\cdot\mathbb{P}(\xi \in I).$$

The assumption $a_1 \geq 1/2$ yields $\frac{1-a_1}{a_1} \leq 1$. Since $\mathrm{Var}(\xi) = \gamma$, we get by Chebyshev's inequality that

$$\mathbb{P}(\xi \in I) = 1 - \mathbb{P}(|\xi - \mathbb{E}\xi| \geq 2) \geq 1 - \frac{1}{4}\mathrm{Var}(\xi) = 1 - \frac{\gamma}{4}.$$

Putting these together and the trivial bound $a_1^{\gamma/2} \leq 1$, we get the lower bound $M(Z) \geq c_\gamma(1-a_1)^{\frac{\gamma-1}{2}}$ with $c_\gamma = \frac{4^\gamma(4-\gamma)}{4^2e^4\Gamma(\gamma)}$.

Combining the two cases together, the lower bound in (8) holds with

$$c_\gamma = \min\left\{(2^{\frac{3-\gamma}{2}}\sqrt{3\gamma})^{-1}, \frac{4^\gamma(4-\gamma)}{4^2e^4\Gamma(\gamma)}\right\} \geq \min\left\{\frac{1}{2^{3/2}\sqrt{3}}, \frac{3\gamma}{4^2e^4}\right\} > 0.003\gamma,$$

since $\frac{1}{\Gamma(\gamma)} \geq \gamma$. $\square$

References

[1] Bartczak, M., Nayar, P., Zwara, S., Sharp variance-entropy comparison for nonnegative Gaussian quadratic forms, *IEEE Trans. Inform. Theory* 67 (2021), no. 12, 7740–7751.

[2] Bhatia, R., Matrix analysis. Graduate Texts in Mathematics, 169. *Springer-Verlag, New York*, 1997.

[3] Bobkov, S. G., Chistyakov, G. P., On concentration functions of random variables. *J. Theoret. Probab.* 28 (2015), no. 3, 976–988.

[4] Bobkov, S. G., Naumov, A. A., Ulyanov, V. V., Two-sided inequalities for the density function's maximum of weighted sum of chi-square variables, preprint (2020) arXiv:2012.10747. In: Recent Developments in Stochastic Methods and Applications. Springer Proceedings in Mathematics & Statistics, 371 (2021), pp. 178–189. Springer, Cham.

[5] Chasapis, G., Gurushankar, K., Tkocz, T., Sharp bounds on p-norms for sums of independent uniform random variables, $0 < p < 1$, preprint (2021), arXiv:2105.14079. To appear in *J. Anal. Math.*

[6] Eskenazis, A., Nayar, P., Tkocz, T., Gaussian mixtures: entropy and geometric inequalities, *Ann. of Prob.* 46(5) 2018, 2908–2945.

[7] Eskenazis, A., Nayar, P., Tkocz, T., Sharp comparison of moments and the log-concave moment problem. *Adv. Math.* 334 (2018), 389–416.

[8] Feller, W., An introduction to probability theory and its applications. Vol. II. Second edition *John Wiley & Sons, Inc., New York-London-Sydney* 1971.

[9] Jameson, G. J. O., A simple proof of Stirling's formula for the gamma function. Math. Gaz. 99 (2015), no. 544, 68–74.

[10] Moriguti, S., A lower bound for a probability moment of any absolutely continuous distribution with finite variance. Ann. Math. Statistics 23 (1952), 286–289.

[11] Nguyen, H. H., Vu, V. H., Small ball probability, inverse theorems, and applications. *Erdös centennial*, 409–463, Bolyai Soc. Math. Stud., 25, *János Bolyai Math. Soc.*, Budapest, 2013.

Carnegie Mellon University; Pittsburgh, PA 15213, USA.

Email address: {gchasapi,salils,ttkocz}@andrew.cmu.edu