Submitted to the Annals of Probability
arXiv: arXiv:1611.04921

GAUSSIAN MIXTURES: ENTROPY AND GEOMETRIC INEQUALITIES

By Alexandros Eskenazis[*,‡], Piotr Nayar[*,†,§] and Tomasz Tkocz[*,‡]

Princeton University[‡] *and University of Pennsylvania*[§]

A symmetric random variable is called a *Gaussian mixture* if it has the same distribution as the product of two independent random variables, one being positive and the other a standard Gaussian random variable. Examples of Gaussian mixtures include random variables with densities proportional to $e^{-|t|^p}$ and symmetric p-stable random variables, where $p \in (0,2]$. We obtain various sharp moment and entropy comparison estimates for weighted sums of independent Gaussian mixtures and investigate extensions of the B-inequality and the Gaussian correlation inequality in the context of Gaussian mixtures. We also obtain a correlation inequality for symmetric geodesically convex sets in the unit sphere equipped with the normalized surface area measure. We then apply these results to derive sharp constants in Khinchine inequalities for vectors uniformly distributed on the unit balls with respect to p-norms and provide short proofs to new and old comparison estimates for geometric parameters of sections and projections of such balls.

1. Introduction. Gaussian random variables and processes have always been of central importance in probability theory and have numerous applications in various areas of mathematics. Recall that the measure γ_n on $\mathbb{R}^n$ with density $\mathrm{d}\gamma_n(x) = (2\pi)^{-n/2} e^{-\sum_{j=1}^n x_j^2/2}\,\mathrm{d}x$ is called the standard Gaussian measure and a random vector distributed according to γ_n is called a standard Gaussian random vector. A centered Gaussian measure on $\mathbb{R}^n$ is defined to be a linear image of standard Gaussian measure. In the past four decades intensive research has been devoted to geometric properties related to Gaussian measures (see, e.g., the survey [28]), which have provided indispensable tools for questions in convex geometry and the local theory of Banach spaces. In many cases, however, it still remains a challenging open problem to determine whether such properties are Gaussian per se or, in

[*]The author was supported in part by the Simons Foundation.

[†]The author was supported in part by NCN grant DEC-2012/05/B/ST1/00412.

MSC 2010 subject classifications: Primary 60E15; secondary 52A20, 52A40, 94A17.

Keywords and phrases: Gaussian measure, Gaussian mixture, Khinchine inequality, entropy, B-inequality, small ball probability, correlation inequalities, extremal sections and projections of ℓ_p balls.

fact, more general.

The main purpose of the present article is to investigate properties of mixtures of Gaussian measures and demonstrate that they are of use to concrete geometric questions.

Definition 1. *A random variable X is called a (centered) Gaussian mixture if there exists a positive random variable Y and a standard Gaussian random variable Z, independent of Y, such that X has the same distribution as the product YZ.*

For example, a random variable X with density of the form

$$f(x) = \sum_{j=1}^{m} p_j \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{x^2}{2\sigma_j^2}},$$

where $p_j, \sigma_j > 0$ are such that $\sum_{j=1}^{m} p_j = 1$, is a Gaussian mixture corresponding to the discrete random variable Y with $\mathbb{P}(Y = \sigma_j) = p_j$. Finite weighted averages of non-centered Gaussian measures are ubiquitous in information theory and theoretical computer science (see, for instance, [16], [1] for relevant results in learning theory) and are often referred in the literature as Gaussian mixtures. In this paper, we shall reserve this term for *centered* Gaussian mixtures in the sense of Definition 1. Observe that Gaussian mixtures are necessarily symmetric and continuous. We shall now discuss a simple analytic characterization of Gaussian mixtures in terms of their probability density functions.

Recall that an infinitely differentiable function $g : (0, \infty) \to \mathbb{R}$ is called completely monotonic if $(-1)^n g^{(n)}(x) \geqslant 0$ for all $x > 0$ and $n \geqslant 0$, where for $n \geqslant 1$ we denote by $g^{(n)}$ the n-th derivative of g and $g^{(0)} = g$. A classical theorem of Bernstein (see, e.g., [17]) asserts that g is completely monotonic if and only if it is the Laplace transform of some measure, i.e. there exists a non-negative Borel measure μ on $[0, \infty)$ such that

$$f(x) = \int_0^\infty e^{-tx} \, \mathrm{d}\mu(t), \quad \text{for every } x > 0. \tag{1}$$

Bernstein's theorem implies the following equivalence.

Theorem 2. *A symmetric random variable X with density f is a Gaussian mixture if and only if the function $x \mapsto f(\sqrt{x})$ is completely monotonic for $x > 0$.*

Theorem 2 will be proven in Section 2. It readily implies that for every $p \in (0,2]$ the random variable with density $c_p e^{-|x|^p}$ is a Gaussian mixture; we denote its law by μ_p and by $\mu_p^n = \mu_p^{\otimes n}$ the corresponding product measure. Another example of Gaussian mixtures are symmetric p-stable random variables, where $p \in (0,2]$ (see Lemma 23 in Section 2). Recall that a symmetric p-stable random variable X is a random variable whose characteristic function is $\mathbb{E}e^{itX} = e^{-c|t|^p}$, for $t \in \mathbb{R}$ and some $c > 0$. Standard symmetric p-stable random variables correspond to $c = 1$. In the consecutive subsections we shall describe our main results on Gaussian mixtures.

1.1. *Sharp Khinchine-type inequalities.* The classical Khinchine inequality asserts that for every $p \in (0,\infty)$ there exist positive constants A_p, B_p such that for every real numbers $a_1, \ldots, a_n$ we have

$$A_p\Big(\sum_{i=1}^n a_i^2\Big)^{1/2} \leqslant \Big(\mathbb{E}\Big|\sum_{i=1}^n a_i\varepsilon_i\Big|^p\Big)^{1/p} \leqslant B_p\Big(\sum_{i=1}^n a_i^2\Big)^{1/2}, \tag{2}$$

where $\varepsilon_1, \ldots, \varepsilon_n \in \{-1,1\}$ are independent symmetric random signs. Whittle discovered the best constants in (2) for $p \geqslant 3$ (see [55]), Szarek treated the case $p = 1$ (see [52]) and finally Haagerup completed this line of research determining the optimal values of A_p, B_p for any $p > 0$ (see [20]).

Following Haagerup's results, sharp Khinchine inequalities for other random variables have also been investigated extensively (see, for example, [30], [5], [41], [27]). In particular, in [30], Latała and Oleszkiewicz treated the case of i.i.d. random variables uniformly distributed on $[-1,1]$ and proved a comparison result in the sense of majorization that we shall now describe.

We say that a vector $a = (a_1, \ldots, a_n)$ is majorized by a vector $b = (b_1, \ldots, b_n)$, denoted $a \preceq b$, if the nonincreasing rearrangements $a_1^* \geqslant \ldots \geqslant a_n^*$ and $b_1^* \geqslant \ldots \geqslant b_n^*$ of the coordinates of a and b, respectively, satisfy the inequalities

$$\sum_{j=1}^k a_j^* \leqslant \sum_{j=1}^k b_j^* \quad \text{for each } k \in \{1, \ldots, n-1\} \quad \text{and} \quad \sum_{j=1}^n a_j = \sum_{j=1}^n b_j.$$

For a general reference on properties and applications of the majorization ordering see [36]. For instance, every vector $(a_1, \ldots, a_n)$ with $a_i \geqslant 0$ and $\sum_{i=1}^n a_i = 1$ satisfies

$$\Big(\frac{1}{n}, \ldots, \frac{1}{n}\Big) \preceq (a_1, \ldots, a_n) \preceq (1, 0, \ldots, 0). \tag{3}$$

A real-valued function which preserves (respectively reverses) the ordering $\preceq$ is called Schur convex (respectively Schur concave). The majorization ordering has many equivalent definitions. For example, it is well-known (see [36]) that a vector $(a_1,\ldots,a_n)$ is majorized by another vector $(b_1,\ldots,b_n)$ if and only if for every continuous convex function $g:\mathbb{R}\to\mathbb{R}$ we have

$$\sum_{i=1}^n g(a_i) \leqslant \sum_{i=1}^n g(b_i). \tag{4}$$

This is equivalent to saying that the uniform probability measure $\mu_a = \frac{1}{n}\sum_{i=1}^n \delta_{a_i}$ is smaller than $\mu_b = \frac{1}{n}\sum_{i=1}^n \delta_{b_i}$ in the convex ordering, i.e. the property that

$$\int_{\mathbb{R}} g(x)\ \mathrm{d}\mu_a(x) \leqslant \int_{\mathbb{R}} g(x)\ \mathrm{d}\mu_b(x)$$

for every convex function $g:\mathbb{R}\to\mathbb{R}$.

The main result of [30] reads as follows. Let $U_1,\ldots,U_n$ be i.i.d. random variables, uniformly distributed on $[-1,1]$. For $(a_1,\ldots,a_n),(b_1,\ldots,b_n)\in\mathbb{R}^n$ and $p\geqslant 2$ we have

$$(a_1^2,\ldots,a_n^2) \preceq (b_1^2,\ldots,b_n^2) \implies \mathbb{E}\Big|\sum_{i=1}^n a_iU_i\Big|^p \geqslant \mathbb{E}\Big|\sum_{i=1}^n b_iU_i\Big|^p \tag{5}$$

and for $p\in[1,2)$ the second inequality is reversed. In particular, combining (3) and (5), for any $p\geqslant 2$ and a unit vector $(a_1,\ldots,a_n)$ we get

$$\mathbb{E}|U_1|^p \leqslant \mathbb{E}\Big|\sum_{i=1}^n a_iU_i\Big|^p \leqslant \mathbb{E}\Big|\frac{U_1+\cdots+U_n}{\sqrt{n}}\Big|^p, \tag{6}$$

whereas for $p\in[1,2)$ the reverse inequalities hold. Inequality (6) along with the central limit theorem implies that the sharp constants in the Khinchine inequality

$$A_p\Big(\mathbb{E}\Big|\sum_{i=1}^n a_iU_i\Big|^2\Big)^{1/2} \leqslant \Big(\mathbb{E}\Big|\sum_{i=1}^n a_iU_i\Big|^p\Big)^{1/p} \leqslant B_p\Big(\mathbb{E}\Big|\sum_{i=1}^n a_iU_i\Big|^2\Big)^{1/2} \tag{7}$$

are precisely

$$A_p = \begin{cases} \gamma_p, & p\in[1,2) \\ \frac{3^{1/2}}{(p+1)^{1/p}}, & p\in[2,\infty) \end{cases} \quad \text{and} \quad B_p = \begin{cases} \frac{3^{1/2}}{(p+1)^{1/p}}, & p\in[1,2) \\ \gamma_p, & p\in[2,\infty) \end{cases}, \tag{8}$$

where $\gamma_p = \sqrt{2}\left(\frac{\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}}\right)^{1/p}$ is the p-th moment of a standard Gaussian random variable.

Theorem 3 below is an analogue of the Schur monotonicity statement (5) for moments of Gaussian mixtures. Recall that for a random variable Y and $p \neq 0$ we denote by $\|Y\|_p = (\mathbb{E}|Y|^p)^{1/p}$ its p-th moment and $\|Y\|_0 = \exp(\mathbb{E}\log|Y|)$. Notice that since a standard Gaussian random variable Z satisfies $\mathbb{E}|Z|^p = \infty$ for every $p \leqslant -1$, a moment comparison result for Gaussian mixtures can only make sense for p-th moments, where $p > -1$.

Theorem 3. *Let X be a Gaussian mixture and $X_1, \ldots, X_n$ be independent copies of X. For two vectors $(a_1, \ldots, a_n), (b_1, \ldots, b_n)$ in $\mathbb{R}^n$ and $p \geqslant 2$ we have*

$$(a_1^2, \ldots, a_n^2) \preceq (b_1^2, \ldots, b_n^2) \implies \Big\| \sum_{i=1}^n a_i X_i \Big\|_p \leqslant \Big\| \sum_{i=1}^n b_i X_i \Big\|_p, \tag{9}$$

whereas for $p \in (-1, 2)$ the second inequality is reversed, provided that $\mathbb{E}|X|^p < \infty$.

The proof of Theorem 3 and the straightforward derivation of sharp constants for the corresponding Khinchine inequalities (Corollary 25) will be provided in Section 3.

Remark 4. After the submission of this paper, we learned from C. Houdré that Theorem 3 had previously appeared in his joint work [4, Proposition 2.6] with R. Averkamp. We are grateful to C. Houdré for providing us this reference.

As an application we derive similar Schur monotonicity properties for vectors uniformly distributed on the unit ball of ℓ_q^n for $q \in (0, 2]$, which were first considered by Barthe, Guédon, Mendelson and Naor in [9]. Recall that for a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $q > 0$ we denote $\|x\|_q = \big(\sum_{i=1}^n |x_i|^q\big)^{1/q}$ and $\|x\|_\infty = \max_{1\leqslant i\leqslant n} |x_i|$. We also write ℓ_q^n for the quasi-normed space $(\mathbb{R}^n, \|\cdot\|_q)$ and $B_q^n = \{x \in \mathbb{R}^n : \|x\|_q \leqslant 1\}$ for its closed unit ball. In [9], the authors discovered a representation for the uniform measure on B_q^n, relating it to the product measures μ_q^n defined after Theorem 2, and used it to determine the sharp constants in Khinchine inequalities on B_q^n up to a constant factor. Using their representation along with Theorem 3 we deduce the following comparison result.

Corollary 5. *Fix $q \in (0, 2]$ and let $X = (X_1, \ldots, X_n)$ be a random vector uniformly distributed on B_q^n. For two vectors $(a_1, \ldots, a_n), (b_1, \ldots, b_n)$*

in $\mathbb{R}^n$ *and* $p \geqslant 2$ *we have*

$$(a_1^2, \ldots, a_n^2) \preceq (b_1^2, \ldots, b_n^2) \implies \Big\| \sum_{i=1}^n a_i X_i \Big\|_p \leqslant \Big\| \sum_{i=1}^n b_i X_i \Big\|_p, \tag{10}$$

whereas for $p \in (-1, 2)$ *the second inequality is reversed.*

The derivation of the sharp constants in the corresponding Khinchine inequality is postponed to Corollary 26. Given Corollary 5 and the result of [30], which corresponds to the unit cube B_∞^n, the following question seems natural.

Question 6. Let $X = (X_1, \ldots, X_n)$ be a random vector uniformly distributed on B_q^n for some $q \in (2, \infty)$. What are the sharp constants in the Khinchine inequalities for X?

It will be evident from the proof of Corollary 5 that Question 6 is equivalent to finding the sharp Khinchine constants for μ_q^n, where $q \in (2, \infty)$. We conjecture that a Schur monotonicity result, identical to the one in (5), is valid.

1.2. *Entropy comparison.* For a random variable X with density function $f : \mathbb{R} \to \mathbb{R}_+$ the Shannon entropy of X is a fundamental quantity in information theory, defined as

$$h(X) = -\int_{\mathbb{R}} f(x) \log f(x) \, \mathrm{d}x = \mathbb{E}[-\log f(X)],$$

provided that the integral exists. Jensen's inequality yields that among random variables with a fixed variance, the Gaussian random variable maximizes the entropy. Moreover, Pinsker's inequality (see, e.g., [18, Theorem 1.1]) asserts that if a random variable X has variance one and G is a standard Gaussian random variable, then the entropy gap $h(G) - h(X)$ dominates the total variation distance between the laws of X and G. Consequently, the entropy can be interpreted as a measure of *closeness to Gaussianity*. The following question seems natural.

Question 7. Fix $n \geqslant 2$ and suppose that $X_1, \ldots, X_n$ are i.i.d. random variables with finite variance. For which unit vectors $(a_1, \ldots, a_n)$ is the entropy of $\sum_{i=1}^n a_i X_i$ maximized?

The constraint $\sum_{i=1}^n a_i^2 = 1$ on $(a_1, \ldots, a_n)$ plainly fixes the variance of the weighted sum $\sum_{i=1}^n a_i X_i$ and the answer would give the corresponding *most Gaussian* weights.

The first result concerning the entropy of weighted sums of i.i.d. random variables was the celebrated entropy power inequality, first stated by Shannon in [49] and rigorously proven by Stam in [51]. An equivalent formulation of the Shannon-Stam inequality (see [34]) reads as follows. For every $\lambda \in [0,1]$ and independent random variables X, Y we have

$$h(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \geqslant \lambda h(X) + (1-\lambda)h(Y), \tag{11}$$

provided that all the entropies exist. It immediately follows from (11) that if $X_1, \ldots, X_n$ are i.i.d. random variables with finite variance and $(a_1, \ldots, a_n)$ is a unit vector, then we have

$$h\Big(\sum_{i=1}^n a_i X_i\Big) \geqslant h(X_1). \tag{12}$$

In other words, the corresponding minimum in Question 7 is achieved at the direction vectors e_i.

Moreover, a deep monotonicity result in the entropic Central Limit Theorem was obtained in the work [3] of Artstein-Avidan, Ball, Barthe and Naor. The authors proved that for any random variable X with finite variance and any $n \geqslant 1$ we have

$$h\Big(\sum_{i=1}^n \frac{1}{\sqrt{n}} X_i\Big) \leqslant h\Big(\sum_{i=1}^{n+1} \frac{1}{\sqrt{n+1}} X_i\Big), \tag{13}$$

where $X_1, X_2, \ldots$ are independent copies of X.

Given inequality (13), a natural guess for Question 7 would be that the vector $\big(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\big)$ is a maximizer for any $n \geqslant 2$ and for any square-integrable random variable X. However, this is not correct in general. In [7, Proposition 2], the authors showed that for a certain symmetric random variable X uniformly distributed on the union of two intervals the Shannon entropy of the weighted sum $\sqrt{\lambda}X_1 + \sqrt{1-\lambda}X_2$ is not maximized at $\lambda = \frac{1}{2}$.

Nonetheless, for Gaussian mixtures it is possible to obtain the comparison for Rényi entropies which confirms the natural guess. Recall that for a random variable X with density $f : \mathbb{R} \to \mathbb{R}_+$ and $\alpha > 0$, $\alpha \neq 1$, the Rényi entropy of order α of X is defined as

$$h_\alpha(X) = \frac{1}{1-\alpha} \log\Big(\int_{\mathbb{R}} f^\alpha(x)\,\mathrm{d}x\Big).$$

Note that if for some $\alpha > 1$ the integral of f^α is finite, then $h_\alpha(X)$ tends to $h(X)$ as $\alpha \to 1^+$ (see [11, Lemma V.3]), which we shall also denote by $h_1(X)$ for convenience.

THEOREM 8. *Let $X_1, \ldots, X_n$ be i.i.d. Gaussian mixtures and $\alpha \geqslant 1$. Then for two vectors $(a_1, \ldots, a_n), (b_1, \ldots, b_n)$ in $\mathbb{R}^n$ we have*

$$(a_1^2, \ldots, a_n^2) \preceq (b_1^2, \ldots, b_n^2) \implies h_\alpha\Big(\sum_{i=1}^n a_i X_i\Big) \geqslant h_\alpha\Big(\sum_{i=1}^n b_i X_i\Big), \tag{14}$$

provided that all the entropies are finite. In particular, for every unit vector $(a_1, \ldots, a_n)$

$$h(X_1) \leqslant h\Big(\sum_{i=1}^n a_i X_i\Big) \leqslant h\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big). \tag{15}$$

Extensions of inequality (15), even for the uniform measure on the cube, appear to be unknown.

QUESTION 9. Let $U_1, \ldots, U_n$ be i.i.d. random variables, each uniformly distributed on $[-1, 1]$. Is it correct that for every unit vector $(a_1, \ldots, a_n)$

$$h\Big(\sum_{i=1}^n a_i U_i\Big) \leqslant h\Big(\frac{U_1 + \cdots + U_n}{\sqrt{n}}\Big) \ ? \tag{16}$$

Geometrically, this would mean that, in the entropy sense, the *most Gaussian* direction of the unit cube B_∞^n is the main diagonal.

REMARK 10. Since for every $n \geqslant 1$,

$$\Big(\frac{1}{n}, \ldots, \frac{1}{n}\Big) \preceq \Big(\frac{1}{n-1}, \ldots, \frac{1}{n-1}, 0\Big),$$

the monotonicity of entropy inequality (13) for Gaussian mixtures is a direct consequence of Theorem 8. Furthermore, the same holds true for all Rényi entropies of order α, where $\alpha \geqslant 1$.

We close this subsection with an intriguing question in the spirit of the well known fact that a Gaussian random variable has maximum entropy among all random variables with a specified variance. Note that Theorem 8 along with

$$(1, 1, 0, \ldots, 0) \succeq \Big(1, \frac{1}{2}, \frac{1}{2}, 0, \ldots, 0\Big) \succeq \cdots \succeq \Big(1, \frac{1}{n}, \ldots, \frac{1}{n}\Big)$$

imply that for every i.i.d. Gaussian mixtures $X_1, X_2, \ldots$ the sequence $h\big(X_1 + \frac{X_2+\cdots+X_{n+1}}{\sqrt{n}}\big)$, $n = 1, 2, \ldots$ is increasing and in particular

$$h(X_1 + X_2) \leqslant h\Big(X_1 + \frac{X_2 + \cdots + X_{n+1}}{\sqrt{n}}\Big).$$

Thus, the following result should not be surprising.

Proposition 11. *Let X_1, X_2 be independent Gaussian mixtures with finite variance. Then*

$$h(X_1 + X_2) \leqslant h(X_1 + G), \tag{17}$$

where G is a Gaussian random variable independent of X_1 having the same variance as X_2.

We pose a question as to whether this is true in general, under the additional assumption that X_1, X_2 are identically distributed.

Question 12. Let X_1, X_2 be i.i.d. continuous random variables with finite variance. Is it true that

$$h(X_1 + X_2) \leqslant h(X_1 + G), \tag{18}$$

where G is a Gaussian random variable independent of X_1 having the same variance as X_2?

The preceding entropy comparison results will be proven in Section 3.

1.3. *Geometric properties of Gaussian mixtures.* Recall that a function $\varphi : \mathbb{R}^n \to \mathbb{R}_+$ is called log-concave if $\varphi = e^{-V}$ for some convex function $V : \mathbb{R}^n \to (-\infty, \infty]$. A measure μ on $\mathbb{R}^n$ is called log-concave if for every Borel sets $A, B \subseteq \mathbb{R}^n$ and $\lambda \in (0, 1)$ we have

$$\mu(\lambda A + (1 - \lambda)B) \geqslant \mu(A)^{\lambda} \mu(B)^{1-\lambda}. \tag{19}$$

A random vector is called log-concave if it is distributed according to a log-concave measure. Two important examples of log-concave measures on $\mathbb{R}^n$ are Gaussian measures and uniform measures supported on convex bodies. The geometry of log-concave measures, in analogy with the asymptotic theory of convex bodies, has been intensively studied and many major results are known (see, for example, the monograph [2]). The Gaussian measure, however, possesses many delicate properties which are either wrong or whose validity is still unknown for other log-concave measures. In what follows, we will explain how to extend, in the context of Gaussian mixtures, two such properties: the B-inequality, proven by Cordero-Erausquin, Fradelizi and Maurey in [15], and the Gaussian correlation inequality, recently proven by Royen in [45].

Choosing the sets A, B in (19) to be dilations of a fixed convex set $K \subseteq \mathbb{R}^n$ we deduce that for every $a, b > 0$ and $\lambda \in (0,1)$

$$\mu\big((\lambda a + (1-\lambda)b)K\big) \geqslant \mu(aK)^{\lambda}\mu(bK)^{1-\lambda}. \tag{20}$$

The (weak) B-inequality provides a substantial strengthening of (20) for Gaussian measure, under an additional symmetry assumption: for any origin symmetric convex set $K \subseteq \mathbb{R}^n$, $a, b > 0$ and $\lambda \in (0,1)$

$$\gamma_n(a^{\lambda}b^{1-\lambda}K) \geqslant \gamma_n(aK)^{\lambda}\gamma_n(bK)^{1-\lambda}, \tag{21}$$

or, in other words, the function $t \mapsto \gamma_n(e^t K)$ is log-concave on $\mathbb{R}$. In fact, in [15] the following strong form of the above inequality was proven.

Theorem 13 (strong B-inequality, [15]). *Let K be a symmetric convex set in $\mathbb{R}^n$. Then, the function*

$$\mathbb{R}^n \ni (t_1, \ldots, t_n) \longmapsto \gamma_n(\Delta(e^{t_1}, \ldots, e^{t_n})K) \tag{22}$$

is log-concave on $\mathbb{R}^n$, where $\Delta(s_1, \ldots, s_n)$ is the diagonal $n \times n$ matrix with entries $s_1, \ldots, s_n$.

The authors also proved that the same conclusion holds for an arbitrary unconditional log-concave measure, provided that the convex set K is unconditional as well (see [15, Section 5] for further details). Furthermore, they asked whether the weak B-inequality (21) holds for any symmetric log-concave measure and symmetric convex set K; this is currently known as the B-conjecture. We note that in [47], Saroglou confirmed the B-conjecture on the plane (the case of uniform measures on convex planar sets had previously been treated in [35]). Our result in this direction is the following theorem.

Theorem 14. *Let $X_1, \ldots, X_n$ be Gaussian mixtures such that X_i has the same distribution as $Y_i Z_i$, where Y_i is positive and Z_i is a standard Gaussian random variable independent of Y_i. Denote by μ_i the law of X_i and by μ the product measure $\mu_1 \otimes \cdots \otimes \mu_n$. If, additionally, $\log Y_i$ is log-concave for each i, then for every symmetric convex set K in $\mathbb{R}^n$ the function*

$$\mathbb{R}^n \ni (t_1, \ldots, t_n) \longmapsto \mu(\Delta(e^{t_1}, \ldots, e^{t_n})K) \tag{23}$$

is log-concave on $\mathbb{R}^n$.

We do not know whether the additional assumption on the Y_i can be omitted, but we verified (Corollary 30) that both the measure with density proportional to $e^{-|t|^p}$ and the symmetric p-stable measure have this property for $p \in (0,1]$, whereas they do not for $p \in (1,2)$. Notice that the corresponding product measures, apart from μ_1^n, are not log-concave. We note that extending the B-inequality to μ_p^n, where $p > 2$, is of importance. For instance, it has been proven by Saroglou [46] that the weak B-inequality for μ_∞^n (that is, the uniform measure on the unit cube B_∞^n) would imply the conjectured logarithmic Brunn-Minkowski inequality (see [14]) in its full generality. The proof of Theorem 14 will be given in Section 4.

An application of the B-inequality for Gaussian measure is a small ball probability estimate due to Latała and Oleszkiewicz [31]. For a symmetric convex set K denote by $r(K)$ its inradius, i.e. the largest $r > 0$ such that $rB_2^n \subseteq K$. In [31], the authors used Theorem 13 along with the Gaussian isoperimetric inequality (see, e.g., [2, Theorem 3.1.9]) to prove that if $K \subseteq \mathbb{R}^n$ is a symmetric convex set with $\gamma_n(K) \leqslant 1/2$, then

$$\gamma_n(tK) \leqslant (2t)^{\frac{r(K)^2}{4}} \gamma_n(K), \quad \text{for every } t \in [0,1]. \tag{24}$$

Using Theorem 14 and an isoperimetric-type estimate of Bobkov and Houdré from [13] we deduce the following corollary.

Corollary 15. *Let K be a symmetric convex set in $\mathbb{R}^n$ such that $\mu_1^n(K) \leqslant 1/2$. Then*

$$\mu_1^n(tK) \leqslant t^{\frac{r(K)}{2\sqrt{6}}} \mu_1^n(K), \quad \textit{for every } t \in [0,1]. \tag{25}$$

Our next result is an extension of the Gaussian correlation inequality, which was recently proven by Royen in [45] (see also [29] for a very clear exposition of Royen's proof and the references therein for the history of the problem).

Theorem 16 (Gaussian correlation inequality, [45]). *For any centered Gaussian measure γ on $\mathbb{R}^n$ and symmetric convex sets K, L in $\mathbb{R}^n$ we have*

$$\gamma(K \cap L) \geqslant \gamma(K)\gamma(L). \tag{26}$$

This inequality admits a straightforward extension to products of laws of Gaussian mixtures.

THEOREM 17. *Let $X_1, \dots, X_n$ be Gaussian mixtures and denote by μ_i the law of X_i. Then, for $\mu = \mu_1 \otimes \cdots \otimes \mu_n$ and any symmetric convex sets K, L in $\mathbb{R}^n$ we have*

$$\mu(K \cap L) \geqslant \mu(K)\mu(L). \tag{27}$$

This theorem implies that the correlation inequality (27) holds for the product measure μ_p^n as well as for all symmetric p-stable laws on $\mathbb{R}^n$, where $p \in (0,2)$ (Corollary 35). In particular, the multivariate Cauchy distribution, which is a rotationally invariant 1-stable distribution on $\mathbb{R}^n$ defined as $\mathrm{d}\mu(x) = c_n(1+\|x\|_2^2)^{-\frac{n+1}{2}}\,\mathrm{d}x$, satisfies the inequality (27). In [39], Memarian proved partial results in this direction and noticed that such inequalities are equivalent to correlation-type inequalities on the unit sphere S^{n-1}. We will recap his argument in Section 5. Let $S_+^{n-1} \subseteq S^{n-1}$ be the open upper hemisphere, i.e. $S_+^{n-1} = S^{n-1} \cap \{x \in \mathbb{R}^n : \ x_n > 0\}$ whose pole is the point $p = (0, \dots, 0, 1)$. A subset $A \subseteq S_+^{n-1}$ is called geodesically convex if for any two points $x, y \in A$ the shortest arc of the great circle joining x, y is contained in A. Furthermore, A is called symmetric (with respect to the pole p) if for any $x \in A$, the point $x^* \neq x$ which lies on the great circle joining x and p and satisfies $d_{S^{n-1}}(x,p) = d_{S^{n-1}}(p, x^*)$, also belongs in A. Here $d_{S^{n-1}}$ denotes the geodesic distance on the sphere.

COROLLARY 18. *Let $S_+^{n-1} \subseteq S^{n-1}$ be the open upper hemisphere. Then for every symmetric geodesically convex sets K, L in S_+^{n-1} we have*

$$|K \cap L| \cdot |S_+^{n-1}| \geqslant |K| \cdot |L|, \tag{28}$$

where $|\cdot|$ denotes the surface area measure on S^{n-1}.

Finally, we want to stress that one cannot expect that all geometric properties of the Gaussian measure will extend mutatis mutandis to Gaussian mixtures. For example, it has been proven by Bobkov and Houdré in [12] that the Gaussian isoperimetric inequality actually characterizes Gaussian measures. Nevertheless, it might be the case that there are many more that admit such an extension.

1.4. *Sections and projections of B_q^n.* The study of quantitative parameters of sections and projections of convex bodies is a classical topic in convex geometry (for example, see the monograph [24]). As a first application, we revisit two well known theorems and reprove them using some relevant Gaussian mixture representations.

Denote by H_1 the hyperplane $(1,0,\ldots,0)^\perp$ and by H_n the hyperplane $(1,\ldots,1)^\perp$. It has been proven by Barthe and Naor in [10] that for any $q\in(2,\infty]$ and any hyperplane $H\subseteq\mathbb{R}^n$ we have

$$|\mathrm{Proj}_{H_1}B_q^n|\leqslant|\mathrm{Proj}_{H}B_q^n|\leqslant|\mathrm{Proj}_{H_n}B_q^n|, \tag{29}$$

where $|\cdot|$ denotes Lebesgue measure. To deduce this, they proved that for any $q\in[1,\infty]$, if $X_1,\ldots,X_n$ are i.i.d. random variables with density

$$f_q(t)=c_q|t|^{\frac{2-q}{q-1}}e^{-|t|^{\frac{q}{q-1}}},\quad t\in\mathbb{R}, \tag{30}$$

then the volume of hyperplane projections of B_q^n can be expressed as

$$|\mathrm{Proj}_{a^\perp}B_q^n|=\alpha_{q,n}\mathbb{E}\Big|\sum_{i=1}^n a_iX_i\Big|, \tag{31}$$

where $a=(a_1,\ldots,a_n)$ is a unit vector and $\alpha_{q,n}$ is a positive constant. It immediately follows from the characterization given in Theorem 2 that for $q\geqslant 2$ the random variables X_i are Gaussian mixtures and thus, from Theorem 3 (with $p=1$), we deduce the following strengthening of (29).

Corollary 19. *Fix $q\in(2,\infty]$. For two unit vectors $a=(a_1,\ldots,a_n), b=(b_1,\ldots,b_n)$ in $\mathbb{R}^n$ we have*

$$(a_1^2,\ldots,a_n^2)\preceq(b_1^2,\ldots,b_n^2)\implies|\mathrm{Proj}_{a^\perp}B_q^n|\geqslant|\mathrm{Proj}_{b^\perp}B_q^n|. \tag{32}$$

We now turn to the dual question for sections. Meyer and Pajor and later Koldobsky (see [40], [23]) proved that for any $q\in(0,2)$ and any hyperplane $H\subseteq\mathbb{R}^n$

$$|B_q^n\cap H_n|\leqslant|B_q^n\cap H|\leqslant|B_q^n\cap H_1|. \tag{33}$$

More precisely, in [40] the authors proved the upper bound of (33) for $q\in[1,2)$ and the lower bound for $q=1$ and posed a conjecture that would imply (33) for any $q\in(0,2)$; this was later confirmed in [23]. The main ingredients in Koldobsky's proof of (33) were a general representation of the volume of hyperplane sections of a convex body in terms of the Fourier transform of the underlying norm and an elegant lemma about symmetric q-stable densities. Using a different approach, we prove the analogue of Corollary 19 for sections.

Corollary 20. *Fix $q\in(0,2)$. For two unit vectors $a=(a_1,\ldots,a_n), b=(b_1,\ldots,b_n)$ in $\mathbb{R}^n$ we have*

$$(a_1^2,\ldots,a_n^2)\preceq(b_1^2,\ldots,b_n^2)\implies|B_q^n\cap a^\perp|\leqslant|B_q^n\cap b^\perp|. \tag{34}$$

In fact, Corollary 20 will follow from a more general comparison of Gaussian parameters of sections which is in the spirit of [9]. For a hyperplane $H \subseteq \mathbb{R}^n$ and a convex body $K \subseteq \mathbb{R}^n$ denote by $\|\cdot\|_{K\cap H}$ the norm on H associated with the convex body $K \cap H$.

THEOREM 21. *Fix $q \in (0,2)$. For a unit vector $\theta \in \mathbb{R}^n$ let G_θ be a standard Gaussian random vector on the hyperplane $\theta^\perp$. Then for every $\lambda > 0$ and unit vectors $a = (a_1, \ldots, a_n), b = (b_1, \ldots, b_n)$ in $\mathbb{R}^n$ we have*

$$(a_1^2, \ldots, a_n^2) \preceq (b_1^2, \ldots, b_n^2) \implies \mathbb{E}e^{-\lambda\|G_a\|^q_{B_q^n\cap a^\perp}} \leqslant \mathbb{E}e^{-\lambda\|G_b\|^q_{B_q^n\cap b^\perp}}. \tag{35}$$

In [9], the authors used a different method to prove that for any $q \in (0,2)$ and $\lambda > 0$ the Gaussian parameters appearing in (35) are maximized when $a = e_1$. As explained there, such inequalities imply the comparison of various other parameters of sections and projections of B_q^n, most notably the volume (Corollary 20) and the mean width. Recall that for a symmetric convex body K in $\mathbb{R}^n$ the support function $h_K : S^{n-1} \to \mathbb{R}_+$ is defined as $h_K(\theta) = \max_{x\in K}\langle x, \theta\rangle$ and the mean width is

$$w(K) = \int_{S^{n-1}} h_K(\theta)\,\mathrm{d}\sigma(\theta),$$

where σ is the rotationally invariant probability measure on the unit sphere S^{n-1}. Exploiting the duality between sections and projections we deduce the following corollary.

COROLLARY 22. *Fix $q \in (2,\infty]$ and let $H \subseteq \mathbb{R}^n$ be a hyperplane. Then*

$$w(\mathrm{Proj}_{H_1} B_q^n) \leqslant w(\mathrm{Proj}_H B_q^n) \leqslant w(\mathrm{Proj}_{H_n} B_q^n). \tag{36}$$

The lower bound in (36) was first obtained in [9], where the authors also proved that for any $q \in (0,2)$ and any hyperplane $H \subseteq \mathbb{R}^n$

$$w(\mathrm{Proj}_H B_q^n) \leqslant w(\mathrm{Proj}_{H_1} B_q^n). \tag{37}$$

Given this result and Corollary 22, what remains to be understood is which hyperplane projections of B_q^n have minimal mean width for $q \in (0,2)$, similarly to the study of volume. We will provide the proof of Theorem 21 and its consequences in Section 6.

2. Proof of Theorem 2 and examples. Here we establish some initial facts about Gaussian mixtures, prove the characterization presented in the introduction and use it to provide relevant examples.

Let X be a Gaussian mixture with the same distribution as YZ, where Y is positive and Z is an independent standard Gaussian random variable; denote by ν the law of Y. Clearly X is symmetric. Furthermore, for a Borel set $A \subseteq \mathbb{R}$ we have

$$\begin{aligned}\mathbb{P}(X \in A) &= \mathbb{P}(YZ \in A) \\ &= \int_0^\infty \mathbb{P}(yZ \in A)\,\mathrm{d}\nu(y) = \int_A \int_0^\infty \frac{1}{\sqrt{2\pi}y} e^{-\frac{x^2}{2y^2}}\,\mathrm{d}\nu(y)\,\mathrm{d}x,\end{aligned} \tag{38}$$

which immediately implies that X has a density

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x^2}{2y^2}} \frac{\mathrm{d}\nu(y)}{y}. \tag{39}$$

We now proceed with the proof of Theorem 2.

Proof of Theorem 2. Let X be a symmetric random variable with density f such that the function $x \mapsto f(\sqrt{x})$ is completely monotonic. By Bernstein's theorem, there exists a non-negative Borel measure μ supported on $[0, \infty)$ such that

$$f(\sqrt{x}) = \int_0^\infty e^{-tx}\,\mathrm{d}\mu(t), \quad \text{for every } x > 0 \tag{40}$$

or, equivalently, $f(x) = \int_0^\infty e^{-tx^2}\,\mathrm{d}\mu(t)$ for every $x \in \mathbb{R}$. Notice that $\mu(\{0\}) = 0$, because otherwise f would not be integrable. Now, for a subset $A \subseteq \mathbb{R}$ we have

$$\begin{aligned}\mathbb{P}(X \in A) &= \int_A \int_0^\infty e^{-tx^2}\,\mathrm{d}\mu(t)\,\mathrm{d}x = \int_0^\infty \int_A e^{-tx^2}\,\mathrm{d}x\,\mathrm{d}\mu(t) \\ &= \int_0^\infty \int_{\sqrt{2t}A} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,\mathrm{d}x\,\sqrt{\frac{\pi}{t}}\,\mathrm{d}\mu(t) \\ &= \int_0^\infty \gamma_n(\sqrt{2t}A)\,\mathrm{d}\nu(t),\end{aligned} \tag{41}$$

where $\mathrm{d}\nu(t) = \sqrt{\frac{\pi}{t}}\,\mathrm{d}\mu(t)$. In particular, choosing $A = \mathbb{R}$, we deduce that ν is a probability measure, supported on $(0, \infty)$. Let V be a random variable distributed according to ν; clearly V is positive almost surely. Define $Y =$

$\frac{1}{\sqrt{2V}}$ and let Z be a standard Gaussian random variable, independent of Y. Then (41) implies that

$$\mathbb{P}(YZ \in A) = \mathbb{P}\left(\frac{1}{\sqrt{2V}} \cdot Z \in A\right) = \int_0^\infty \gamma_n(\sqrt{2t}A)\,\mathrm{d}\nu(t) = \mathbb{P}(X \in A),$$

that is, X has the same distribution as the product YZ. The converse implication readily follows from (39) and Bernstein's theorem after a change of variables. □

Before applying Theorem 2 we first provide some examples of completely monotonic functions. Differentiation shows that the functions $e^{-\alpha x}, x^{-\alpha}$ and $(1+x)^{-\alpha}$, where $\alpha > 0$, are completely monotonic on $(0,\infty)$ and a straightforward induction proves that the same holds for e^{-x^β}, where $\beta \in (0,1]$. The same argument implies that if g is a completely monotonic function on $(0,\infty)$ and h is positive and has a completely monotonic derivative on $(0,\infty)$, then $g \circ h$ is also completely monotonic on $(0,\infty)$. Moreover, one can easily see that products of completely monotonic functions themselves are completely monotonic.

Combining the last example with Theorem 2, we get that for every $p \in (0,2]$ the random variable with density proportional to $e^{-|t|^p}$ is a Gaussian mixture. Recall that we denote by μ_p the probability measure with density $c_p e^{-|t|^p}$, $p > 0$, where $c_p = (2\Gamma(1+1/p))^{-1}$, and $\mu_p^n = \mu_p^{\otimes n}$. Furthermore, it is a classical fact that symmetric p-stable random variables, where $p \in (0,2]$, are Gaussian mixtures. For these measures we can describe the positive factor in their Gaussian mixture representation. Recall that a positive random variable W with Laplace transform $\mathbb{E}e^{-tW} = e^{-ct^\alpha}$, where $\alpha \in (0,1)$ and $c > 0$, is called a positive α-stable random variable. Standard positive α-stable random variables correspond to $c = 1$; we denote their density by g_α.

LEMMA 23. *Fix $p \in (0,2)$ and let Z be a standard Gaussian random variable.*

(i) If $V_{p/2}$ has density proportional to $t^{-1/2}g_{p/2}(t)$ and is independent of Z, then $(2V_{p/2})^{-1/2}Z$ has density $c_p e^{-|t|^p}$.

(ii) If $W_{p/2}$ is a standard positive $p/2$-stable random variable and is independent of Z, then $(2W_{p/2})^{1/2}Z$ is a standard symmetric p-stable random variable.

PROOF. To show (i), we shall decompose a symmetric random variable with density $c_p e^{-|x|^p}$ into a product of two independent random variables: a positive one and a standard Gaussian. To this end, denote by μ the measure in the representation (40) written for the density $c_p e^{-|x|^p}$, that is

$$c_p e^{-x^{p/2}} = \int_0^\infty e^{-tx}\, \mathrm{d}\mu(t), \quad x > 0.$$

Therefore, the Laplace transform of $c_p^{-1}\mu$ is $e^{-x^{p/2}}$, which implies that $c_p^{-1}\mu$ is a standard positive $p/2$-stable measure with density $g_{p/2}$. Now, an inspection of the proof of Theorem 2, reveals that the positive factor Y in the Gaussian mixture representation is $Y = (2V)^{-1/2}$, where V has law $\sqrt{\frac{\pi}{t}}\, \mathrm{d}\mu(t)$, so in this case the density of V is indeed proportional to $t^{-1/2} g_{p/2}(t)$, as required.

On the other hand, (ii) is a straightforward characteristic function computation. Using the independence of $W_{p/2}$ and Z we get

$$\mathbb{E} e^{i\sqrt{2} t W_{p/2}^{1/2} Z} = \mathbb{E}_{W_{p/2}} \mathbb{E}_Z e^{i\sqrt{2} t W_{p/2}^{1/2} Z} = \mathbb{E} e^{-t^2 W_{p/2}} = e^{-t^p},$$

which concludes the proof of the lemma. □

Lemma 23 will be useful in Section 4. For instance, a direct computation shows that these Gaussian mixture representations have the following explicit forms when $p = 1$.

(i) Let $\mathcal{E}$ be an exponential random variable (that is, a random variable with density $e^{-t}\mathbf{1}_{t>0}$) and Z a standard Gaussian random variable, independent of $\mathcal{E}$. Then the product $\sqrt{2\mathcal{E}}Z$ has density $\frac{1}{2}e^{-|t|}$, $t \in \mathbb{R}$ (symmetric exponential density).

(ii) Let Z_1, Z_2 be independent standard Gaussian random variables. Then the quotient $Z_1/|Z_2|$ is distributed according to the Cauchy distribution with density $\frac{1}{\pi(1+x^2)}$, which is the symmetric 1-stable distribution.

Remark 24. It was noted in [10, p. 223] that for an infinitely differentiable integrable function $f : (0, \infty) \to \mathbb{R}$, the function $x \mapsto f(\sqrt{x})$ is completely monotonic if and only if $x \mapsto \widehat{f}(\sqrt{x})$ is completely monotonic, where $\widehat{f}$ is the Fourier transform of f. Applying this to the density $c_p e^{-|t|^p}$ and then using Theorem 2 yields that symmetric p-stable random variables are Gaussian mixtures, as was also proven above.

3. Moment and entropy comparison. For the proofs of this section, we will use an elementary result of Marshall and Proschan from [37] which reads as follows. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a convex function, symmetric under permutations of its n arguments. Let $X_1, \ldots, X_n$ be interchangeable random variables, that is, random variables whose joint distribution is invariant under permutations of its coordinates. Then for two vectors $(a_1,\ldots,a_n),(b_1,\ldots,b_n)\in\mathbb{R}^n$ we have

$$(a_1,\ldots,a_n)\preceq(b_1,\ldots,b_n) \implies \mathbb{E}\phi(a_1X_1,\ldots,a_nX_n)\leqslant\mathbb{E}\phi(b_1X_1,\ldots,b_nX_n) \tag{42}$$

or, in other words, the function $\mathbb{R}^n\ni(a_1,\ldots,a_n)\mapsto\mathbb{E}\phi(a_1X_1,\ldots,a_nX_n)$ is Schur convex. If ϕ is concave, then the second inequality in (42) is reversed, i.e. the function above is Schur concave. This result follows directly from the fact that a convex (respectively concave) function that is symmetric under permutations of its arguments is Schur convex (respectively concave), which, in turn, is a consequence of the following simple property. If $a=(a_1,\ldots,a_n),b=(b_1,\ldots,b_n)\in\mathbb{R}^n$ then

$$a\preceq b \iff a\in\mathrm{conv}\big\{(b_{\sigma(1)},\ldots,b_{\sigma(n)}):\ \sigma\text{ is a permutation of }\{1,\ldots,n\}\big\},$$

where $\mathrm{conv}(A)$ denotes the convex hull of a set $A\subseteq\mathbb{R}^n$ (for details, see [36]).

We start with the comparison of moments of Gaussian mixtures.

Proof of Theorem 3. Fix $p>-1$, $p\neq0$. Let X be a Gaussian mixture and $X_1,\ldots,X_n$ be independent copies of X. Since each X_i is a Gaussian mixture, there exist i.i.d. positive random variables $Y_1,\ldots,Y_n$ and independent standard Gaussian random variables $Z_1,\ldots,Z_n$ such that X_i has the same distribution as the product Y_iZ_i. For $a_1,\ldots,a_n\in\mathbb{R}$ the joint independence of the Y_i,Z_j implies that

$$\mathbb{E}\Big|\sum_{i=1}^n a_iX_i\Big|^p=\mathbb{E}\Big|\sum_{i=1}^n a_iY_iZ_i\Big|^p=\mathbb{E}\Big|\Big(\sum_{i=1}^n a_i^2Y_i^2\Big)^{1/2}Z\Big|^p=\gamma_p^p\cdot\mathbb{E}\Big|\sum_{i=1}^n a_i^2Y_i^2\Big|^{p/2},$$

where Z is a standard Gaussian random variable independent of all the Y_i and $\gamma_p=(\mathbb{E}|Z|^p)^{1/p}$. The conclusion now follows directly from Marshall and Proschan's result (42) since $t\mapsto t^{p/2}$ is convex for $p\in(-1,0)\cup[2,\infty)$ and concave for $p\in(0,2)$. Notice that when the exponent $1/p$ is negative, the resulting norm becomes Schur concave. The result for $p=0$ is proven similarly. □

The derivation of sharp constants in the corresponding Khinchine inequalities is now straightforward.

Corollary 25. *Let X be a Gaussian mixture and $X_1,\ldots,X_n$ be independent copies of X. Then, for every $p\in(-1,\infty)$ and $a_1,\ldots,a_n$ in $\mathbb{R}$ we* *have*

$$A_p\Big\|\sum_{i=1}^n a_iX_i\Big\|_2 \leqslant \Big\|\sum_{i=1}^n a_iX_i\Big\|_p \leqslant B_p\Big\|\sum_{i=1}^n a_iX_i\Big\|_2, \tag{43}$$

where

$$A_p = \begin{cases} \frac{\|X\|_p}{\|X\|_2}, & p\in(-1,2) \\ \gamma_p, & p\in[2,\infty) \end{cases} \quad \textit{and} \quad B_p = \begin{cases} \gamma_p, & p\in(-1,2) \\ \frac{\|X\|_p}{\|X\|_2}, & p\in[2,\infty) \end{cases}, \tag{44}$$

provided that all the moments exist. Here $\gamma_p = \sqrt{2}\left(\frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}}\right)^{1/p}$ *is the* p*-th moment of a standard Gaussian random variable. These constants are sharp.*

Proof. We can clearly assume that $(a_1,\ldots,a_n)$ is a unit vector. We will prove the statement for $p\geqslant 2$; the case $p\in(-1,2)$ is identical. The Schur convexity statement of Theorem 3 along with (3) implies that

$$\Big\|\frac{X_1+\cdots+X_n}{\sqrt{n}}\Big\|_p \leqslant \Big\|\sum_{i=1}^n a_iX_i\Big\|_p \leqslant \|X_1\|_p. \tag{45}$$

Applying this for $a_1=\cdots=a_{n-1}=(n-1)^{-1/2}$ and $a_n=0$, where $n\geqslant 2$, shows that the quantity on the left-hand side is decreasing in n and the central limit theorem implies that

$$\gamma_p\|X\|_2 \leqslant \Big\|\sum_{i=1}^n a_iX_i\Big\|_p \leqslant \|X\|_p,$$

which is equivalent to

$$\gamma_p\Big\|\sum_{i=1}^n a_iX_i\Big\|_2 \leqslant \Big\|\sum_{i=1}^n a_iX_i\Big\|_p \leqslant \frac{\|X\|_p}{\|X\|_2}\Big\|\sum_{i=1}^n a_iX_i\Big\|_2.$$

The sharpness of the constants is evident. □

For the proof of Corollary 5 we need to exploit two results about the geometry of B_q^n which are probabilistic in nature. Let $Y_1,\ldots,Y_n$ be i.i.d. random variables distributed according to μ_q and write $Y=(Y_1,\ldots,Y_n)$.

We denote by S the random variable $\left(\sum_{i=1}^n |Y_i|^q\right)^{1/q}$. As explained in the introduction, the main ingredient of the proof of Corollary 5 is a representation for the uniform measure on B_q^n discovered in [9] that reads as follows. Let $\mathcal{E}$ be an exponential random variable (that is, the density of $\mathcal{E}$ is $e^{-t}\mathbf{1}_{t>0}$) independent of the Y_i. Then the random vector

$$\Big(\frac{Y_1}{(S^q+\mathcal{E})^{1/q}},\ldots,\frac{Y_n}{(S^q+\mathcal{E})^{1/q}}\Big)$$

is uniformly distributed on B_q^n. We will also need a result of Schechtman and Zinn from [48], also independently proven by Rachev and Rüschendorf in [44], which asserts that the random variables S and $\frac{Y}{S}$ are independent.

Proof of Corollary 5. Recall that $X=(X_1,\ldots,X_n)$ is a random vector uniformly distributed on B_q^n and let $Y_1,\ldots,Y_n, S$ and $\mathcal{E}$ be as above. For the reader's convenience we repeat the following computation from [9]. Using the representation described before and the independence of S and $\frac{Y}{S}$ we get

$$\mathbb{E}\Big|\sum_{i=1}^n a_iX_i\Big|^p = \mathbb{E}\Big|\frac{1}{(S^q+\mathcal{E})^{1/q}}\sum_{i=1}^n a_iY_i\Big|^p = \mathbb{E}\Big|\frac{S}{(S^q+\mathcal{E})^{1/q}}\Big|^p\mathbb{E}\Big|\sum_{i=1}^n a_i\frac{Y_i}{S}\Big|^p.$$

Then, again by independence, $\mathbb{E}\big|\sum_{i=1}^n a_i\frac{Y_i}{S}\big|^p\mathbb{E}|S|^p = \mathbb{E}\big|\sum_{i=1}^n a_iY_i\big|^p$ and thus

$$\begin{aligned}\mathbb{E}\Big|\sum_{i=1}^n a_iX_i\Big|^p &= \frac{1}{\mathbb{E}|S|^p}\mathbb{E}\Big|\frac{S}{(S^q+\mathcal{E})^{1/q}}\Big|^p\mathbb{E}\Big|\sum_{i=1}^n a_iY_i\Big|^p \\ &= c(p,q,n)\mathbb{E}\Big|\sum_{i=1}^n a_iY_i\Big|^p,\end{aligned} \tag{46}$$

where $c(p,q,n)>0$ is independent of the vector $(a_1,\ldots,a_n)$. In other words, the moments of linear functionals applied to the vector X are proportional to the moments of the same linear functionals applied to Y. In view of Theorem 3 and of the fact that $Y_1,\ldots,Y_n$ are i.i.d. Gaussian mixtures, this property readily implies Corollary 5. $\square$

Similarly to Corollary 25, it is straightforward to deduce the sharp constants for Khinchine inequalities on B_q^n.

COROLLARY 26. *Fix $q\in(0,2]$ and let $X=(X_1,\ldots,X_n)$ be a random vector, uniformly distributed on B_q^n. Then, for every $p\in(-1,\infty)$ and $a_1,\ldots,a_n$ in $\mathbb{R}$ we have*

$$A_p\Big\|\sum_{i=1}^n a_iX_i\Big\|_2 \leqslant \Big\|\sum_{i=1}^n a_iX_i\Big\|_p \leqslant B_p\Big\|\sum_{i=1}^n a_iX_i\Big\|_2, \tag{47}$$

where

$$A_p = \begin{cases} \frac{\|X_1\|_p}{\|X_1\|_2}, & p \in (-1,2) \\ \gamma_p, & p \in [2,\infty) \end{cases} \quad \text{and} \quad B_p = \begin{cases} \gamma_p, & p \in (-1,2) \\ \frac{\|X_1\|_p}{\|X_1\|_2}, & p \in [2,\infty) \end{cases} \tag{48}$$

and for $r > -1$

$$\|X_1\|_r = \left(\frac{B\Big(\frac{r+1}{q}, \frac{n+q-1}{q}\Big)}{B\Big(\frac{1}{q}, \frac{n+q-1}{q}\Big)} \right)^{1/r}, \tag{49}$$

These constants are sharp.

Proof. The derivation of (48) is identical to the one in the proof of Corollary 25. To deduce (49), notice that X_1 has density $f(x) = c_{q,n}(1-|x|^q)^{\frac{n-1}{q}}\mathbf{1}_{|x|\leqslant 1}$ where

$$c_{q,n}^{-1} = 2\int_0^1 (1-x^q)^{\frac{n-1}{q}}\,\mathrm{d}x = \frac{2}{q}B\Big(\frac{1}{q}, \frac{n+q-1}{q}\Big).$$

Thus, for every $r > 0$,

$$\|X_1\|_r = \Big(2c_{q,n}\int_0^1 x^r(1-x^q)^{\frac{n-1}{q}}\,\mathrm{d}x\Big)^{1/r} = \left(\frac{B\Big(\frac{r+1}{q}, \frac{n+q-1}{q}\Big)}{B\Big(\frac{1}{q}, \frac{n+q-1}{q}\Big)} \right)^{1/r},$$

which completes the proof. $\square$

We now turn to comparison of entropy.

Proof of Theorem 8. Let X be a Gaussian mixture and $X_1, \ldots, X_n$ independent copies of X. There exist i.i.d. positive random variables $Y_1, \ldots, Y_n$ and independent standard Gaussian random variables $Z_1, \ldots, Z_n$ such that X_i has the same distribution as the product Y_iZ_i. For a vector $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{R}^n$ denote by X_θ the random variable $\sum_{i=1}^n \theta_i X_i$ and by f_θ the density of X_θ. Since X_θ is itself a Gaussian mixture, Theorem 2 implies that the function $x \mapsto f_\theta(\sqrt{x})$ is completely monotonic. Consequently, there exists a measure μ_θ on $[0,\infty)$ so that

$$f_\theta(\sqrt{x}) = \int_0^\infty e^{-tx}\,\mathrm{d}\mu_\theta(t), \quad \text{for every } x > 0.$$

It now immediately follows from Hölder's inequality that for $x, y > 0$ and $\lambda \in (0,1)$ we have

$$\begin{aligned} f_\theta(\sqrt{\lambda x + (1-\lambda)y}) &= \int_0^\infty (e^{-tx})^\lambda (e^{-ty})^{1-\lambda} \,\mathrm{d}\mu_\theta(t) \\ &\leqslant \Big(\int_0^\infty e^{-tx} \,\mathrm{d}\mu_\theta(t)\Big)^\lambda \Big(\int_0^\infty e^{-ty} \,\mathrm{d}\mu_\theta(t)\Big)^{1-\lambda} \\ &= f_\theta(\sqrt{x})^\lambda f_\theta(\sqrt{y})^{1-\lambda} \end{aligned}$$

or, in other words, the function $\varphi_\theta(x) = -\log f_\theta(\sqrt{x})$ is concave.

Let $a = (a_1, \ldots, a_n), b = (b_1, \ldots, b_n) \in \mathbb{R}^n$ be such that $(a_1^2, \ldots, a_n^2) \preceq (b_1^2, \ldots, b_n^2)$. We first consider the case of Shannon entropy, i.e. $\alpha = 1$. Jensen's inequality implies the following well known variational formula

$$\begin{aligned} h(X_b) &= \mathbb{E}[-\log f_b(X_b)] \\ &= \min\Big\{\mathbb{E}[-\log g(X_b)] : \ g : \mathbb{R} \to \mathbb{R}_+ \text{ is a density function}\Big\}. \end{aligned} \tag{50}$$

Thus, using (50) for $g = f_a$ we get

$$\begin{aligned} h(X_b) &\leqslant \mathbb{E}[-\log f_a(X_b)] = \mathbb{E}\Big[-\log f_a\Big(\sum_{i=1}^n b_i Y_i Z_i\Big)\Big] \\ &= \mathbb{E}\Big[-\log f_a\Big(\Big(\sum_{i=1}^n b_i^2 Y_i^2\Big)^{1/2} Z\Big)\Big] = \mathbb{E}_Z \mathbb{E}_Y \varphi_a\Big(\sum_{i=1}^n b_i^2 Y_i^2 Z^2\Big), \end{aligned} \tag{51}$$

where in the last equality we used the fact that Z is independent of the Y_i. Now, since $(a_1^2, \ldots, a_n^2)$ is majorized by $(b_1^2, \ldots, b_n^2)$, the concavity of φ_a along with Marshall and Proschan's result (42) imply that

$$\mathbb{E}_Y \varphi_a\Big(\sum_{i=1}^n b_i^2 Y_i^2 Z^2\Big) \leqslant \mathbb{E}_Y \varphi_a\Big(\sum_{i=1}^n a_i^2 Y_i^2 Z^2\Big)$$

which, after averaging over Z, gives

$$h(X_b) \leqslant \mathbb{E}\varphi_a\Big(\sum_{i=1}^n a_i^2 Y_i^2 Z^2\Big) = \mathbb{E}[-\log f_a(X_a)] = h(X_a).$$

For the Rényi entropy of order α, where $\alpha > 1$, we need to prove that

$$\int_{\mathbb{R}} f_a^\alpha(x) \,\mathrm{d}x \leqslant \int_{\mathbb{R}} f_b^\alpha(x) \,\mathrm{d}x. \tag{52}$$

Notice that, as before, we can write

$$\int_{\mathbb{R}} f_a^\alpha(x)\,\mathrm{d}x = \mathbb{E} f_a^{\alpha-1}(X_a) = \mathbb{E}_Z \mathbb{E}_Y f_a^{\alpha-1}\Big(\Big(\sum_{i=1}^n a_i^2 Y_i^2\Big)^{1/2} Z\Big). \tag{53}$$

The concavity of φ_a implies that, since $\alpha > 1$, the function $x \mapsto f_a^{\alpha-1}(\sqrt{x}) = e^{(1-\alpha)\varphi_a(x)}$ is convex and thus from (42) we get

$$\mathbb{E}_Y f_a^{\alpha-1}\Big(\Big(\sum_{i=1}^n a_i^2 Y_i^2\Big)^{1/2} Z\Big) \leqslant \mathbb{E}_Y f_a^{\alpha-1}\Big(\Big(\sum_{i=1}^n b_i^2 Y_i^2\Big)^{1/2} Z\Big)$$

which, after integrating with respect to Z, gives

$$\begin{aligned} \int_{\mathbb{R}} f_a^\alpha(x)\,\mathrm{d}x &\leqslant \mathbb{E} f_a^{\alpha-1}\Big(\Big(\sum_{i=1}^n b_i^2 Y_i^2\Big)^{1/2} Z\Big) \\ &= \mathbb{E} f_a^{\alpha-1}(X_b) = \int_{\mathbb{R}} f_a^{\alpha-1}(x) f_b(x)\,\mathrm{d}x. \end{aligned} \tag{54}$$

Finally, Hölder's inequality yields

$$\int_{\mathbb{R}} f_a^{\alpha-1}(x) f_b(x)\,\mathrm{d}x \leqslant \Big(\int_{\mathbb{R}} f_a^\alpha(x)\,\mathrm{d}x\Big)^{\frac{\alpha-1}{\alpha}} \Big(\int_{\mathbb{R}} f_b^\alpha(x)\,\mathrm{d}x\Big)^{\frac{1}{\alpha}}. \tag{55}$$

Combining (54) and (55) readily implies (52), i.e. the comparison $h_\alpha(X_a) \geqslant h_\alpha(X_b)$. □

Remark 27. We note that a result of similar nature was proven in the work [56] of Yu, who showed that for every i.i.d. symmetric log-concave random variables $X_1, \ldots, X_n$ the function $(a_1, \ldots, a_n) \mapsto h\big(\sum_{i=1}^n a_i X_i\big)$ is Schur convex on $\mathbb{R}^n$. In particular, for every vector $(a_1, \ldots, a_n) \in \mathbb{R}^n$ such that $\sum_{i=1}^n |a_i| = 1$ we have

$$h\Big(\frac{1}{n}\sum_{i=1}^n X_i\Big) \leqslant h\Big(\sum_{i=1}^n a_i X_i\Big) \leqslant h(X_1). \tag{56}$$

The main actors in Yu's argument are the same: the variational principle for entropy (50) and Marshall and Proschan's comparison result (42) (the log-concavity assumption is paired up with the linear constraint on the coefficients).

Finally, we proceed with the proof of Proposition 11.

Proof of Proposition 11. Let X_1, X_2 be independent Gaussian mixtures such that X_i has the same distribution as the product $Y_i Z_i$, for some independent positive random variables Y_1, Y_2 and independent standard Gaussian random variables Z_1, Z_2. Let G be a centered Gaussian random variable independent of X_1 with the same variance as X_2. Notice that $X_1 + X_2$ has the same distribution as $(Y_1^2 + Y_2^2)^{1/2} Z$, whereas $X_1 + G$ has the same distribution as $(Y_1^2 + \mathbb{E}Y_2^2)^{1/2} Z$, where Z is a standard Gaussian random variable independent of the Y_i. Denote by f the density of $X_1 + X_2$ and by g the density of $X_1 + G$. Using the variational formula for entropy (50) we get

$$\begin{aligned} h(X_1 + X_2) &= \mathbb{E}[-\log f(X_1 + X_2)] \leqslant \mathbb{E}[-\log g(X_1 + X_2)] \\ &= \mathbb{E}_{(Y_1, Z)} \mathbb{E}_{Y_2}[-\log g((Y_1^2 + Y_2^2)^{1/2} Z)]. \end{aligned}$$

Since $X_1 + G$ is also a Gaussian mixture, as remarked in the proof of Theorem 8, the function $x \mapsto -\log g(\sqrt{x})$ is concave and thus

$$\mathbb{E}_{Y_2}[-\log g((Y_1^2 + Y_2^2)^{1/2} Z)] \leqslant -\log g((Y_1^2 + \mathbb{E}Y_2^2)^{1/2} Z).$$

Combining the above we deduce that

$$\begin{aligned} h(X_1 + X_2) &\leqslant \mathbb{E}[-\log g((Y_1^2 + \mathbb{E}Y_2^2)^{1/2} Z)] \\ &= \mathbb{E}[-\log g(X_1 + G)] = h(X_1 + G), \end{aligned}$$

which concludes the proof. □

Remark 28. In light of Proposition 11, it could seem that the assumption that X_1, X_2 are identically distributed in Question 12 is redundant. However, this is not the case. Let X_1, X_2 be independent symmetric random variables such that X_1 has a smooth density $f : \mathbb{R} \to \mathbb{R}_+$ and let G be an independent Gaussian random variable with the same variance as X_2. A straightforward differentiation shows that the inequality

$$h(X_1 + \varepsilon X_2) \leqslant h(X_1 + \varepsilon G)$$

as $\varepsilon \to 0^+$ is equivalent to the comparison of the fourth order Taylor coefficients of these expressions, namely

$$\mathbb{E}X_2^4 \int_{\mathbb{R}} f^{(4)}(x) \log f(x) \, \mathrm{d}x \geqslant \mathbb{E}G^4 \int_{\mathbb{R}} f^{(4)}(x) \log f(x) \, \mathrm{d}x.$$

However, this inequality can easily be seen to be wrong, e.g. by taking X_1 to have density function $f(x) = \frac{x^2}{\sqrt{2\pi}} e^{-x^2/2}$ and X_2 to be uniformly distributed on a symmetric interval.

4. The B-inequality. We start by establishing a straightforward representation for products of laws of Gaussian mixtures. Let $X_1, \ldots, X_n$ be independent Gaussian mixtures (not necessarily identically distributed) so that X_i has the same distribution as the product $Y_i Z_i$, where $Y_1, \ldots, Y_n$ are independent positive random variables and $Z_1, \ldots, Z_n$ are independent standard Gaussian random variables. Denote by ν_i the law of Y_i, by μ_i the law of X_i and by ν, μ the product measures $\nu_1 \otimes \cdots \otimes \nu_n$ and $\mu_1 \otimes \cdots \otimes \mu_n$ respectively. Then, for a Borel set $A \subseteq \mathbb{R}^n$ we have

$$\begin{aligned} \mu(A) &= \mathbb{P}((X_1, \ldots, X_n) \in A) = \mathbb{P}((Y_1 Z_1, \ldots, Y_n Z_n) \in A) \\ &= \int_0^\infty \cdots \int_0^\infty \mathbb{P}((y_1 Z_1, \ldots, y_n Z_n) \in A)\, \mathrm{d}\nu_1(y_1) \cdots \mathrm{d}\nu_n(y_n) \\ &= \int_{(0,\infty)^n} \gamma_n(\Delta(y_1, \ldots, y_n)^{-1} A)\, \mathrm{d}\nu(y_1, \ldots, y_n), \end{aligned} \tag{57}$$

where $\Delta(y_1, \ldots, y_n)$ is the diagonal matrix with entries $y_1, \ldots, y_n$. In other words, μ is an average of centered Gaussian measures on $\mathbb{R}^n$. We now proceed with the proof of the B-inequality for Gaussian mixtures.

Proof of Theorem 14. Let $X_1, \ldots, X_n$ be as in the statement of the theorem and denote by f_i the density of Y_i. Clearly, the log-concavity of the random variable $\log Y_i$ is equivalent to the log-concavity of the function $s \mapsto f_i(e^{-s})$ on $\mathbb{R}$. Let $K \subseteq \mathbb{R}^n$ be a symmetric convex set and $(t_1, \ldots, t_n) \in \mathbb{R}^n$. Then, by (57) and the change of variables $y_i = e^{-s_i}$ we have

(58)

$$\begin{aligned} \mu(\Delta(e^{t_1}, \ldots, e^{t_n})K) &= \int_{(0,\infty)^n} \gamma_n(\Delta(y_1^{-1} e^{t_1}, \ldots, y_n^{-1} e^{t_n})K) \prod_{i=1}^n f_i(y_i)\, \mathrm{d}y \\ &= \int_{\mathbb{R}^n} \gamma_n(\Delta(e^{s_1+t_1}, \ldots, e^{s_n+t_n})K) \prod_{i=1}^n f_i(e^{-s_i}) e^{-\sum_{i=1}^n s_i}\, \mathrm{d}s. \end{aligned}$$

The B-inequality for Gaussian measure (Theorem 13) immediately implies that the function

$$\mathbb{R}^n \times \mathbb{R}^n \ni (s,t) \longmapsto \gamma_n(\Delta(e^{s_1+t_1}, \ldots, e^{s_n+t_n})K)$$

is log-concave on $\mathbb{R}^n \times \mathbb{R}^n$. Consequently, the integrand in (58) is a log-concave function of $(s,t) \in \mathbb{R}^n \times \mathbb{R}^n$ as a product of log-concave functions. The result now follows from the Prékopa-Leindler inequality (see, e.g., [2, Theorem 1.4.1]) which implies that marginals of log-concave functions are log-concave (see also [19, Theorem 3.15]). $\square$

REMARK 29. An inspection of the proof of Theorem 14 shows that the same argument also yields the B-inequality for rotationally invariant measures of the form $\mathrm{d}\mu(x) = f(\|x\|_2)\,\mathrm{d}x$, where f is proportional to the density of a Gaussian mixture that satisfies the assumption of Theorem 14.

Checking whether a particular Gaussian mixture X satisfies the assumption of Theorem 14 might be non-trivial, since one has to know the distribution of the positive factor Y occurring in its representation. However, by Lemma 23, we know this factor for random variables with densities proportional to $e^{-|t|^p}$ and for symmetric p-stable random variables, where $p \in (0,2)$. This allows us to determine the values of $p \in (0,2)$ for which the assumption is satisfied, for each of these random variables.

To this end, denote, as before, by g_α the density of a standard positive α-stable random variable, $\alpha \in (0,1)$. Recall that the positive factor in the representation of a standard symmetric p-stable random variable is $(2W_{p/2})^{1/2}$, where $W_{p/2}$ is a standard positive $p/2$-stable random variable. Thus, the assumption of Theorem 14 is equivalent to the log-concavity of the function $s \mapsto g_{p/2}(e^{-s})$ on $\mathbb{R}$. On the other hand, the corresponding factor in the representation of the random variable with density $c_p e^{-|t|^p}$ is of the form $(2V_{p/2})^{-1/2}$ where $V_{p/2}$ has density proportional to $t^{-1/2}g_{p/2}(t)$. Therefore, the corresponding assumption in this case is again equivalent to the log-concavity of $s \mapsto g_{p/2}(e^{-s})$ on $\mathbb{R}$, since the remaining factor $e^{s/2}$ is log-affine. If X is a random variable with density $g : \mathbb{R} \to \mathbb{R}_+$, the log-concavity of $s \mapsto g(e^{-s})$ is referred in the literature as multiplicative strong unimodality of X. The multiplicative strong unimodality of positive α-stable distributions has been studied by Simon in [50], who proved that such a random variable has this property if and only if $\alpha \leqslant 1/2$. Combining this with the above observations and Theorem 14 we deduce the following.

COROLLARY 30. *For every $p \in (0,1]$ the product measure on $\mathbb{R}^n$ with density proportional to $e^{-\|x\|_p^p}$ and the symmetric p-stable product measure on $\mathbb{R}^n$ satisfy the B-inequality for every symmetric convex set $K \subseteq \mathbb{R}^n$.*

We now turn to the proof of the small ball estimate for the symmetric exponential measure (Corollary 15) described in the introduction. The argument is very similar to the one in [31].

Proof of Corollary 15. Let $K \subseteq \mathbb{R}^n$ be a symmetric convex set such that $\mu_1^n(K) \leqslant 1/2$ and we denote by $r = r(K)$ the inradius of K. For a set $A \subseteq \mathbb{R}^n$ and $h > 0$ we also denote by A_h the h-enlargement of A, that is,

$A_h = A + hB_2^n$. Notice that for $s \in (0,1)$ we have $(sK) \cap (K^c)_{(1-s)r} = \varnothing$, where K^c is the complement of K, and thus

$$\mu_1^n(sK) \leqslant 1 - \mu_1^n((K^c)_{(1-s)r}). \tag{59}$$

Now, choose $u \geqslant 0$ such that $\mu_1^n(K) = \mu_1((u,\infty))$ or, equivalently, $\mu_1^n(K^c) = \mu_1((-u,\infty))$. Bobkov and Houdré proved in [13] that if $A \subseteq \mathbb{R}^n$ is a Borel set and $x \in \mathbb{R}$ is such that $\mu_1^n(A) = \mu_1((x,\infty))$, then for every $h > 0$ we have

$$\mu_1^n(A_h) \geqslant \mu_1\Big(\Big(x - \frac{h}{2\sqrt{6}}, \infty\Big)\Big). \tag{60}$$

Combining (59) and (60) we get

$$\begin{aligned}\mu_1^n(sK) &\leqslant 1 - \mu_1\Big(\Big(-u - \frac{(1-s)r}{2\sqrt{6}}, \infty\Big)\Big) \\ &= \mu_1\Big(\Big(u + \frac{(1-s)r}{2\sqrt{6}}, \infty\Big)\Big) = e^{\frac{s-1}{2\sqrt{6}}r(K)}\mu_1^n(K).\end{aligned} \tag{61}$$

For $0 < t \leqslant s \leqslant 1$ we can write $s = t^{\frac{\log s}{\log t}}$ and the B-inequality for μ_1^n implies that

$$\mu_1^n(tK)^{\frac{\log s}{\log t}} \mu_1^n(K)^{1-\frac{\log s}{\log t}} \leqslant \mu_1^n(sK),$$

or equivalently

$$\frac{\mu_1^n(tK)}{\mu_1^n(K)} \leqslant \left(\frac{\mu_1^n(sK)}{\mu_1^n(K)}\right)^{\frac{\log t}{\log s}}, \tag{62}$$

which, in view of (61), gives the estimate

$$\mu_1^n(tK) \leqslant e^{\frac{s-1}{2\sqrt{6}}\cdot\frac{\log t}{\log s}r(K)}\mu_1^n(K) = t^{\frac{r(K)}{2\sqrt{6}}\cdot\frac{s-1}{\log s}}\mu_1^n(K).$$

Taking the limit $s \to 1^-$ we finally deduce that

$$\mu_1^n(tK) \leqslant t^{\frac{r(K)}{2\sqrt{6}}}\mu_1^n(K),$$

for every $t \in [0,1]$, which concludes the proof. □

REMARK 31. In [42], Paouris and Valettas proved a different small ball probability estimate for the symmetric exponential measure and any unconditional convex body K in terms of the parameter $\beta(K) = \mathrm{Var}\|W\|_K / m(K)^2$, where W is distributed according to μ_1^n and $m(K)$ is the median of $\|\cdot\|_K$

with respect to μ_1^n. Their result is in the spirit of the work [22] of Klartag and Vershynin. In the follow-up paper [43], they showed that a similar estimate holds for every unconditional log-concave measure and unconditional convex body K with a worse dependence on $\beta(K)$. In the particular case of the symmetric exponential measure the unconditionality assumption in the suboptimal estimate from [43] can be omitted, because of Corollary 30.

We would like to remark that Theorem 14 combined with a result of Marsiglietti, [38, Proposition 3.1], immediately implies the following corollary.

COROLLARY 32. *Let μ be as in Theorem 14. Then, for every symmetric convex set $K \subseteq \mathbb{R}^n$ the function $t \mapsto \mu(tK)$ is $\frac{1}{n}$-concave for $t > 0$, that is*

$$\mu\big((\lambda t + (1-\lambda)s)K\big)^{1/n} \geqslant \lambda\mu(tK)^{1/n} + (1-\lambda)\mu(sK)^{1/n}, \tag{63}$$

for every $t, s > 0$ and $\lambda \in (0,1)$.

5. Correlation inequalities. To prove the correlation inequality for Gaussian mixtures (Theorem 17) we will use Royen's Gaussian correlation inequality (Theorem 16), along with a simple lemma for the standard Gaussian measure. Recall that we write $\Delta(y) = \Delta(y_1, \ldots, y_n)$ for the diagonal $n \times n$ matrix with diagonal $y = (y_1, \ldots, y_n)$.

LEMMA 33. *Let K be a symmetric convex set in $\mathbb{R}^n$. Then the function $t \mapsto \gamma_n\big(\Delta(t, 1, \ldots, 1)K\big)$ is nondecreasing for $t > 0$.*

PROOF. It clearly suffices to consider the case when K has nonempty interior. We will prove that the function $\psi(x) = \log \gamma_n\big(\Delta(e^x, 1, \ldots, 1)K\big)$ is nondecreasing on the real line. By virtue of the B-inequality for the standard Gaussian measure (Theorem 13), ψ is concave. To verify that ψ is nondecreasing, it is enough to prove that $\lim_{x\to\infty} \psi(x) > -\infty$. Take $\delta > 0$ such that $[-\delta, \delta]^n \subseteq K$. For every real number x we have

$$\psi(x) \geqslant \log \gamma_n([-e^x\delta, e^x\delta] \times [-\delta, \delta]^{n-1}),$$

which, for $x \to \infty$, gives

$$\lim_{x\to\infty} \psi(x) \geqslant \log \gamma_n(\mathbb{R} \times [-\delta, \delta]^{n-1}) > -\infty.$$

This concludes the proof of the lemma. $\square$

Proof of Theorem 17. Let μ be a product of laws of Gaussian mixtures. According to (57) for every Borel set $A \subseteq \mathbb{R}^n$ we have

$$\mu(A) = \int_{(0,\infty)^n} \gamma_n(\Delta(y)^{-1}A)\,\mathrm{d}\nu_1(y_1)\cdots\mathrm{d}\nu_n(y_n),$$

for some probability measures $\nu_1,\ldots,\nu_n$ on $(0,\infty)$. Let $K, L \subseteq \mathbb{R}^n$ be symmetric convex sets. The Gaussian correlation inequality yields

$$\begin{aligned}\mu(K\cap L) &= \int_{(0,\infty)^n} \gamma_n(\Delta(y)^{-1}K \cap \Delta(y)^{-1}L)\,\mathrm{d}\nu_1(y_1)\cdots\mathrm{d}\nu_n(y_n)\\ &\geqslant \int_{(0,\infty)^n} \gamma_n(\Delta(y)^{-1}K)\gamma_n(\Delta(y)^{-1}L)\,\mathrm{d}\nu_1(y_1)\cdots\mathrm{d}\nu_n(y_n).\end{aligned} \tag{64}$$

Fix $y_1,\ldots,y_{n-1} > 0$. By Lemma 33, the functions $y_n \mapsto \gamma_n(\Delta(y)^{-1}K)$ and $y_n \mapsto \gamma_n(\Delta(y)^{-1}L)$ are nonincreasing on $(0,\infty)$. Consequently, combining (64) and Chebyshev's integral inequality (see, e.g., [21, p. 168]) for the probability measure ν_n, we get

$$\begin{aligned}\mu(K\cap L) \geqslant \int_{(0,\infty)^{n-1}} &\Big(\int_0^\infty \gamma_n(\Delta(y)^{-1}K)\,\mathrm{d}\nu_n(y_n)\Big)\times\\ &\times\Big(\int_0^\infty \gamma_n(\Delta(y)^{-1}L)\,\mathrm{d}\nu_n(y_n)\Big)\,\mathrm{d}\nu_1(y_1)\cdots\mathrm{d}\nu_{n-1}(y_{n-1}).\end{aligned}$$

After iteratively applying Chebyshev's inequality to $\nu_1,\ldots,\nu_{n-1}$ we finally deduce that

$$\begin{aligned}\mu(K\cap L) \geqslant &\Big(\int_{(0,\infty)^n} \gamma_n(\Delta(y)^{-1}K)\,\mathrm{d}\nu_1(y_1)\cdots\mathrm{d}\nu_n(y_n)\Big)\times\\ &\times\Big(\int_{(0,\infty)^n} \gamma_n(\Delta(y)^{-1}L)\,\mathrm{d}\nu_1(y_1)\cdots\mathrm{d}\nu_n(y_n)\Big) = \mu(K)\mu(L),\end{aligned}$$

which is the correlation inequality (27). □

Remark 34. Similarly to the B-inequality, an inspection of the proof of Theorem 17 reveals that the same argument also gives the correlation inequality for rotationally invariant probability measures of the form $\mathrm{d}\mu(x) = f(\|x\|_2)\,\mathrm{d}x$, where f is proportional to the density of a Gaussian mixture.

Recall that a function $f : \mathbb{R}^n \to \mathbb{R}_+$ is called quasiconcave if for any $t \geqslant 0$ the set $A_t = \{x \in \mathbb{R}^n : \; f(x) \geqslant t\}$ is convex. Writing

$$f(x) = \int_0^\infty \mathbf{1}_{A_t}(x)\,\mathrm{d}t, \quad x \in \mathbb{R},$$

one can immediately see that if a measure μ satisfies the correlation inequality (27) for any symmetric convex sets $K, L \subseteq \mathbb{R}^n$ then for every symmetric quasiconcave functions $f, g : \mathbb{R}^n \to \mathbb{R}_+$ we have

$$\int_{\mathbb{R}^n} f(x)g(x)\,\mathrm{d}\mu(x) \geqslant \int_{\mathbb{R}^n} f(x)\,\mathrm{d}\mu(x) \cdot \int_{\mathbb{R}^n} g(x)\,\mathrm{d}\mu(x). \tag{65}$$

Correlation inequalities of the form (65) were treated by Koldobsky and Montgomery-Smith in [26] for another class of functions when μ is a general symmetric stable measure on $\mathbb{R}^n$. Recall that the law μ of a random vector X in $\mathbb{R}^n$ is called a symmetric p-stable measure if every marginal $\langle X, a\rangle$, $a \in \mathbb{R}^n$, is a symmetric p-stable random variable. It is a well known fact (see, e.g., [54, p. 312]) that symmetric p-stable random vectors $X = (X_1, \ldots, X_n)$ in $\mathbb{R}^n$ are in one-to-one correspondence with finite measures m_X on the unit sphere S^{n-1} such that

$$\mathbb{E}\exp\Big(i\sum_{j=1}^n a_jX_j\Big) = \exp\Big(-\int_{S^{n-1}}\Big|\sum_{j=1}^n a_jx_j\Big|^p\,\mathrm{d}m_X(x)\Big), \tag{66}$$

for every $a_1, \ldots, a_n \in \mathbb{R}$. We will argue that the correlation inequality (27) holds for the law μ of any symmetric p-stable random vector X in $\mathbb{R}^n$. Assume first that the corresponding measure m_X on S^{n-1} has a finite support, namely $\mathrm{supp}(m_X) = \{y_1, \ldots, y_\ell\}$, and let Y be a standard ℓ-dimensional symmetric p-stable random vector with independent coordinates. In this case, one can find $\theta_1, \ldots, \theta_n \in \mathbb{R}^\ell$ such that X_j has the same distribution as $\langle Y, \theta_j\rangle$ or, in other words, X is a linear image of Y and the correlation inequality (27) immediately follows. For a general measure m_X on S^{n-1} there exists a sequence of finitely supported measures m_ℓ that converges to m_X in the weak* topology (e.g. by the Krein-Milman theorem) which means, by (66), that the corresponding p-stable random vectors X_ℓ converge to X in distribution. Note that to prove the correlation inequality (27) for a symmetric p-stable measure μ on $\mathbb{R}^n$, it suffices to consider the case when $K, L \subseteq \mathbb{R}^n$ are convex polytopes, which are sets whose boundaries are contained in a finite union of affine hyperplanes. However, any affine hyperplane is of μ-measure zero, since the one-dimensional marginals of μ are p-stable, thus continuous. Therefore, the convergence in distribution concludes the proof of the following corollary.

Corollary 35. *Let μ be a symmetric p-stable measure on $\mathbb{R}^n$. Then for every symmetric convex sets $K, L \subseteq \mathbb{R}^n$ we have*

$$\mu(K \cap L) \geqslant \mu(K)\mu(L).$$

This corollary implies inequalities of the form (65), analogous to the ones proven in [26]. It also implies that the multivariate Cauchy distribution, defined as $\mathrm{d}\mu(x) = c_n(1+\|x\|_2^2)^{-\frac{n+1}{2}}\,\mathrm{d}x$, satisfies the correlation inequality (27). Notice that this also follows from Remark 34. In [39], the author showed that this is actually equivalent to Corollary 18. We reproduce his argument below.

Proof of Corollary 18. Consider the hyperplane $\mathbb{R}^{n-1} \equiv \mathbb{R}^{n-1} \times \{0\} \subseteq \mathbb{R}^n$ and let $S \subseteq \mathbb{R}^n$ be the sphere of radius 1 centered at $e_n = (0,\ldots,0,1)$. Denote by S_+ the open lower hemisphere of S, i.e. $S_+ = \{x \in S : \ x_n < 1\}$, and define a bijection $q : S_+ \to \mathbb{R}^{n-1}$ by the formula

$$q(x) = \text{the point of } \mathbb{R}^{n-1} \text{ which lies on the line joining } x \text{ to } e_n. \tag{67}$$

One can easily check that closed arcs of great circles on S_+ are mapped to line segments on $\mathbb{R}^{n-1}$ and vice versa, which immediately implies that geodesically convex sets in S_+ are in one-to-one correspondence with convex sets in $\mathbb{R}^{n-1}$. Moreover, since $q(0) = 0$, symmetry in $\mathbb{R}^{n-1}$ agrees with geodesic symmetry in S_+. Denoting by μ the push-forward under q of the normalized surface area measure on S_+, we get that for every $r > 0$, μ satisfies the identity

$$\mu(rB_2^{n-1}) = \frac{|B_{S^{n-1}}(\arctan r)|}{|S_+|},$$

where $B_{S^{n-1}}(\theta)$ is a spherical cap of radius θ on S^{n-1}. A simple computation for the volume of spherical caps along with the rotational invariance of μ shows that μ is precisely the law of the multivariate Cauchy distribution on the hyperplane $\mathbb{R}^{n-1}$. Therefore, for two symmetric geodesically convex sets $K, L \subseteq S_+$, the multivariate Cauchy correlation inequality for the symmetric convex sets $q(K), q(L) \subseteq \mathbb{R}^{n-1}$ implies that

$$\frac{|K\cap L|}{|S_+|} = \mu(q(K\cap L)) = \mu(q(K)\cap q(L)) \geqslant \mu(q(K))\cdot\mu(q(L)) = \frac{|K|}{|S_+|}\cdot\frac{|L|}{|S_+|},$$

which completes the proof of the corollary. $\square$

Remark 36. It is a straightforward consequence of Theorem 17 that the product probability measure μ_p^n with density $c_p^n e^{-\|x\|_p^p}$ satisfies the correlation inequality (27) for every $p \in (0,2]$ and $n \geqslant 1$. It turns out that this is the exact range of $p > 0$ for which this property holds. To see this, take $\delta > 0$ and consider the symmetric strips

$$K_\delta = \{(x,y) \in \mathbb{R}^2 : \ |x-y| \leqslant \delta\} \quad \text{and} \quad L_\delta = \{(x,y) \in \mathbb{R}^2 : \ |x+y| \leqslant \delta\}$$

on the plane. We will show that $\mu_p^2(K_\delta \cap L_\delta) < \mu_p^2(K_\delta)\mu_p^2(L_\delta)$ for $p > 2$ and small enough $\delta > 0$. Indeed, a straightforward differentiation yields that the Taylor expansions of these two quantities around $\delta = 0$ are

$$\mu_p^2(K_\delta \cap L_\delta) = 4c_p^2\delta^2 + o(\delta^3) \ \text{ and } \ \mu_p^2(K_\delta)\mu_p^2(L_\delta) = 4c_p^2 2^{1-\frac{2}{p}}\delta^2 + o(\delta^3)$$

and since $1 < 2^{1-\frac{2}{p}}$ for $p > 2$, the correlation inequality (27) cannot hold for small enough $\delta > 0$. A computation along the same lines together with Remark 34 prove that a similar behavior is exhibited by the rotationally invariant probability measures with densities proportional to $e^{-\|x\|_2^p}$: they satisfy (27) if and only if $p \in (0, 2]$.

Remark 37. After the submission of this paper, we learned from J. Zinn about the works [32] and [33] of Lewis and Pritchard on measures supporting correlation inequalities (27). In [32], the authors proved, among other things, that the uniform probability measure on any convex body K in $\mathbb{R}^n$ does not support a correlation inequality, even though we now know that this holds for the uniform measure on a hemisphere, where usual convexity and symmetry are replaced by geodesic convexity and symmetry (Corollary 18). In [33], they showed that any rotationally invariant measure μ on $\mathbb{R}^n$ which supports a correlation inequality must satisfy

$$\int_{\mathbb{R}^n} e^{a\|x\|_2^2}\, \mathrm{d}\mu(x) = \infty$$

for some constant $a > 0$. The computation mentioned in the last sentence of Remark 36 is also a consequence of their result. We are grateful to J. Zinn for providing us these references.

6. Sections and projections of B_q^n revisited. In this section we derive the comparison results for geometric parameters of hyperplane sections and projections of the balls B_q^n described in the introduction. First, let us explain how the comparison of the aforementioned Gaussian parameters (Theorem 21) implies the comparison of volume (Corollary 20) and mean width (Corollary 22), following [9].

Proof of Corollaries 20 and 22. Fix $q \in (0, 2)$ and let $a = (a_1, \ldots, a_n), b = (b_1, \ldots, b_n) \in \mathbb{R}^n$ be unit vectors such that $(a_1^2, \ldots, a_n^2) \preceq (b_1^2, \ldots, b_n^2)$. Recall that G_a, G_b are standard Gaussian random vectors on the hyperplanes $a^\perp$ and $b^\perp$ respectively. According to Theorem 21, for every $\lambda > 0$ we have

$$\mathbb{E}e^{-\lambda\|G_a\|^q_{B_q^n \cap a^\perp}} \leqslant \mathbb{E}e^{-\lambda\|G_b\|^q_{B_q^n \cap b^\perp}}.$$

Integrating this inequality with respect to λ and any measure μ on $(0,\infty)$ we deduce that

$$\mathbb{E}\int_0^\infty e^{-\lambda\|G_a\|^q_{B_q^n\cap a^\perp}}\,\mathrm{d}\mu(\lambda) \leqslant \mathbb{E}\int_0^\infty e^{-\lambda\|G_b\|^q_{B_q^n\cap b^\perp}}\,\mathrm{d}\mu(\lambda),$$

which, by Bernstein's theorem, is equivalent to the validity of the inequality

$$\mathbb{E}g(\|G_a\|^q_{B_q^n\cap a^\perp}) \leqslant \mathbb{E}g(\|G_b\|^q_{B_q^n\cap b^\perp}) \tag{68}$$

for every completely monotonic function $g:(0,\infty)\to\mathbb{R}$. In particular, choosing $g(s)=s^{-\alpha/q}$, we get that

$$\mathbb{E}\|G_a\|^{-\alpha}_{B_q^n\cap a^\perp} \leqslant \mathbb{E}\|G_b\|^{-\alpha}_{B_q^n\cap b^\perp},$$

provided that $0<\alpha<n-1$ so that the integrals are finite. Integration in polar coordinates now shows that for every $0<\alpha<n-1$ we have

$$\int_{S(a^\perp)}\|\theta\|^{-\alpha}_{B_q^n\cap a^\perp}\,\mathrm{d}\sigma_a(\theta) \leqslant \int_{S(b^\perp)}\|\theta\|^{-\alpha}_{B_q^n\cap b^\perp}\,\mathrm{d}\sigma_b(\theta), \tag{69}$$

where σ_a,σ_b are the rotationally invariant probability measures on the unit spheres $S(a^\perp),S(b^\perp)$ of the hyperplanes $a^\perp$ and $b^\perp$, respectively. Letting $\alpha\to n-1$ in (69) now implies that

$$\int_{S(a^\perp)}\|\theta\|^{-n+1}_{B_q^n\cap a^\perp}\,\mathrm{d}\sigma_a(\theta) \leqslant \int_{S(b^\perp)}\|\theta\|^{-n+1}_{B_q^n\cap b^\perp}\,\mathrm{d}\sigma_b(\theta). \tag{70}$$

However, for every symmetric convex body K in $\mathbb{R}^m$, the radius of K in the direction $\theta\in S^{m-1}$ is $\rho_K(\theta)=\sup\{t>0:\ t\theta\in K\}=\|\theta\|_K^{-1}$ and integration in polar coordinates gives

$$|K| = \int_{S^{m-1}}\int_0^{\|\theta\|_K^{-1}} r^{m-1}\,\mathrm{d}r\,\mathrm{d}\theta = \frac{1}{m}\int_{S^{m-1}}\|\theta\|_K^{-m}\,\mathrm{d}\theta,$$

where we denote by $\mathrm{d}\theta$ integration with respect to the usual Lebesgue measure on S^{m-1}. Equivalently, after rescaling we get

$$\int_{S^{m-1}}\|\theta\|_K^{-m}\,\mathrm{d}\sigma(\theta) = \frac{|K|}{|B_2^m|}, \tag{71}$$

which, combined with (70) gives that $|B_q^n\cap a^\perp|\leqslant|B_q^n\cap b^\perp|$ and Corollary 20 follows.

Assume now that $q \in [1,2)$. Applying (68) to $g(s) = e^{-\lambda s^\beta}$, where $\beta \in (0,1]$ and $\lambda > 0$, we have

$$\mathbb{E}e^{-\lambda \|G_a\|^{\beta q}_{B_q^n \cap a^\perp}} \leqslant \mathbb{E}e^{-\lambda \|G_b\|^{\beta q}_{B_q^n \cap b^\perp}}.$$

Since both sides, as functions of $\lambda > 0$, are equal at $\lambda = 0$ we deduce that their derivatives at $\lambda = 0$ also satisfy the same inequality, that is

$$\mathbb{E}\|G_b\|^{\beta q}_{B_q^n \cap b^\perp} \leqslant \mathbb{E}\|G_a\|^{\beta q}_{B_q^n \cap a^\perp},$$

for every $\beta \in (0,1]$. Choosing $\beta = 1/q$ and integrating in polar coordinates yields

$$\int_{S(b^\perp)} \|\theta\|_{B_q^n \cap b^\perp} \, \mathrm{d}\sigma_b(\theta) \leqslant \int_{S(a^\perp)} \|\theta\|_{B_q^n \cap a^\perp} \, \mathrm{d}\sigma_a(\theta). \tag{72}$$

Recall that for a symmetric convex body K in $\mathbb{R}^m$, the polar body K^o of K is defined to be $K^o = \{x \in \mathbb{R}^m : \ \langle x, y\rangle \leqslant 1 \text{ for every } y \in K\}$ and if $\|\cdot\|_{K^o}$ is the norm associated with K^o then $\|\theta\|_{K^o} = h_K(\theta)$ for every $\theta \in S^{n-1}$. Moreover, recall the standard polarity relation (see, e.g., [53, Proposition 2.6]) between sections and projections, namely that for every convex body K on $\mathbb{R}^m$ and every hyperplane H,

$$K \cap H = (\mathrm{Proj}_H(K^o))^o. \tag{73}$$

Combining (72) with (73), we deduce that if $q^* > 2$ is such that $\frac{1}{q} + \frac{1}{q^*} = 1$, then

$$w(\mathrm{Proj}_{b^\perp}(B^n_{q^*})) \leqslant w(\mathrm{Proj}_{a^\perp}(B^n_{q^*})). \tag{74}$$

In particular, for every hyperplane $H \subseteq \mathbb{R}^n$ we obtain

$$w(\mathrm{Proj}_{H_1}(B^n_{q^*})) \leqslant w(\mathrm{Proj}_{H}(B^n_{q^*})) \leqslant w(\mathrm{Proj}_{H_n}(B^n_{q^*})),$$

where $H_1 = (1,0,\ldots,0)^\perp$ and $H_n = (1,\ldots,1)^\perp$. This concludes the proof of Corollary 22. $\square$

We finally proceed with the proof of Theorem 21.

Proof of Theorem 21. Fix $q \in (0,2)$. For a hyperplane $H = a^\perp$, where $a = (a_1,\ldots,a_n) \in \mathbb{R}^n$ is a unit vector, let G_a be a standard Gaussian random vector on $a^\perp$ and denote by $H(\varepsilon)$ the set

$$H(\varepsilon) = \{x \in \mathbb{R}^n : \ |\langle x, a\rangle| < \varepsilon\}. \tag{75}$$

To proceed, we will need a representation from [9, Lemma 14] for the Laplace transforms of $\|G_a\|^q_{B_q^n\cap H}$ that reads as follows. For every $\lambda>0$ there exist constants $\alpha(q,\lambda),\beta(q,\lambda)>0$ and $c(q,\lambda,n)>0$ such that for every hyperplane $H=a^\perp$, $a=(a_1,\dots,a_n)\in S^{n-1}$, we have

$$\mathbb{E}e^{-\lambda\|G_a\|^q_{B_q^n\cap H}}=c(q,\lambda,n)\lim_{\varepsilon\to0^+}\frac{1}{2\varepsilon}\mu^n_{q,\lambda}(H(\varepsilon)),\tag{76}$$

where the probability measure $\mu_{q,\lambda}$ on $\mathbb{R}$ is of the form

$$\mathrm{d}\mu_{q,\lambda}(t)=e^{-\alpha(q,\lambda)|t|^q-\beta(q,\lambda)t^2}\,\mathrm{d}t\tag{77}$$

and $\mu^n_{q,\lambda}=\mu^{\otimes n}_{q,\lambda}$. An immediate application of Theorem 2 yields that $\mu_{q,\lambda}$ is the law of a Gaussian mixture. Thus, by (57) there exists a probability measure $\nu=\nu(q,\lambda)$ on $(0,\infty)$ such that if $A\subseteq\mathbb{R}^n$ is a Borel set, then

$$\mu^n_{q,\lambda}(A)=\int_{(0,\infty)^n}\gamma_n(\Delta(y)^{-1}A)\,\mathrm{d}\nu_n(y),$$

where $\nu_n=\nu^{\otimes n}$. Notice that for the symmetric strip (75) we have

$$\Delta(y)^{-1}H(\varepsilon)=\Big\{x\in\mathbb{R}^n:\Big|\sum_{j=1}^n a_jy_jx_j\Big|<\varepsilon\Big\},$$

that is, $\Delta(y)^{-1}H(\varepsilon)$ is also a symmetric strip of width $\big(\sum_{j=1}^n a_j^2y_j^2\big)^{-1/2}\varepsilon$. Consequently, the rotational invariance of the Gaussian measure implies that

$$\mu^n_{q,\lambda}(H(\varepsilon))=2\int_{(0,\infty)^n}\Psi\Big(\Big(\sum_{j=1}^n a_j^2y_j^2\Big)^{-1/2}\varepsilon\Big)\,\mathrm{d}\nu_n(y),\tag{78}$$

where $\Psi(s)=\frac{1}{\sqrt{2\pi}}\int_0^s e^{-x^2/2}\,\mathrm{d}x$. Combining (76) and (78) we deduce that

$$\begin{aligned}c(q,\lambda,n)^{-1}\cdot\mathbb{E}e^{-\lambda\|G_a\|^q_{B_q^n\cap H}}&=\lim_{\varepsilon\to0^+}\frac{1}{\varepsilon}\int_{(0,\infty)^n}\Psi\Big(\Big(\sum_{j=1}^n a_j^2y_j^2\Big)^{-1/2}\varepsilon\Big)\,\mathrm{d}\nu_n(y)\\&=\int_{(0,\infty)^n}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0}\Psi\Big(\Big(\sum_{j=1}^n a_j^2y_j^2\Big)^{-1/2}\varepsilon\Big)\,\mathrm{d}\nu_n(y)\\&=\frac{1}{\sqrt{2\pi}}\int_{(0,\infty)^n}\Big(\sum_{j=1}^n a_j^2y_j^2\Big)^{-1/2}\,\mathrm{d}\nu_n(y)\\&=\frac{1}{\sqrt{2\pi}}\mathbb{E}\Big(\sum_{j=1}^n a_j^2Y_j^2\Big)^{-1/2},\end{aligned}$$

where $Y_1, \ldots, Y_n$ are i.i.d. random variables distributed according to ν. To verify the assumptions of the dominated convergence theorem for the swap of the limit and integration in the second equality, it suffices to check that $\big(\sum_{j=1}^n a_j^2 y_j^2\big)^{-1/2} \in L_1(\nu_n)$, since $\Psi(s) \leqslant \frac{s}{\sqrt{2\pi}}$ for $s > 0$. This immediately follows by Fatou's lemma, that is

$$\begin{aligned}\int_{(0,\infty)^n} \Big(\sum_{j=1}^n a_j^2 y_j^2\Big)^{-1/2} \mathrm{d}\nu_n(y) &\leqslant \sqrt{2\pi} \liminf_{\varepsilon\to 0^+} \frac{1}{2\varepsilon} \mu_{q,\lambda}^n(H(\varepsilon)) \\ &= \sqrt{2\pi} c(q,\lambda,n)^{-1} \mathbb{E} e^{-\lambda \|G_a\|_{B_q^n\cap H}^q} < \infty.\end{aligned}$$

Now, since $t \mapsto t^{-1/2}$ is a convex function on $(0,\infty)$ and $Y_1, \ldots, Y_n$ are i.i.d. random variables, Marshall and Proschan's result (42) implies the comparison (35), as required. □

We note that the crucial identity (78) can also be proven in purely probabilistic terms. Let $X_1, \ldots, X_n$ be i.i.d. random variables distributed according to $\mu_{q,\lambda}$ and take i.i.d. positive random variables $Y_1, \ldots, Y_n$ and standard Gaussian random variables $Z_1, \ldots, Z_n$ such that X_i has the same distribution as the product $Y_i Z_i$. Then we have

$$\begin{aligned}\mu_{q,\lambda}^n(H(\varepsilon)) &= \mathbb{E}_Y \mathbb{P}_Z\Big(\Big|\sum_{j=1}^n a_j Y_j Z_j\Big| < \varepsilon\Big) = \mathbb{E}_Y \mathbb{P}_Z\Big(|Z| \Big(\sum_{j=1}^n a_j^2 Y_j^2\Big)^{1/2} < \varepsilon\Big) \\ &= \mathbb{E}_Y \mathbb{P}_Z\Big(|Z| < \Big(\sum_{j=1}^n a_j^2 Y_j^2\Big)^{-1/2} \varepsilon\Big) = 2\mathbb{E}\Big[\Psi\Big(\Big(\sum_{j=1}^n a_j^2 Y_j^2\Big)^{-1/2}\varepsilon\Big)\Big],\end{aligned}$$

where Z is a standard Gaussian random variable, independent of the Y_i.

Remark 38. A similar approach also yields a direct proof of Corollary 20. The crucial ingredient in this case would be an identity from [40] and [8] instead of (76). It is proven there that for every $q \in (0,\infty)$, there exists a constant $c(q,n) > 0$ such that if $H \subseteq \mathbb{R}^n$ is any hyperplane and $H(\varepsilon)$ is defined by (75), then

$$|B_q^n \cap H| = c(q,n) \lim_{\varepsilon\to 0^+} \frac{1}{2\varepsilon} \mu_q^n(H(\varepsilon)), \tag{79}$$

where μ_q^n is the measure on $\mathbb{R}^n$ with density proportional to $e^{-\|x\|_q^q}$. Since this measure is also a product of laws of i.i.d. Gaussian mixtures the preceding argument works identically.

In [25], Koldobsky and Zymonopoulou investigated extremal volumes of sections of the complex ℓ_q-balls $B_q^n(\mathbb{C})$, which can also be treated by the approach presented above. From now on we will adopt the obvious identification of $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ without further ado. We will denote by $\langle\cdot,\cdot\rangle$ the standard Hermitian inner product on $\mathbb{C}^n$ and for a vector $\zeta\in\mathbb{C}^n$ we will write $\zeta^\perp$ for the complex hyperplane orthogonal to ζ. Recall that for a vector $z=(x_1,y_1,\ldots,x_n,y_n)\in\mathbb{R}^{2n}$ we denote

$$\|z\|_{\ell_q^n(\mathbb{C})}=\Big(\sum_{j=1}^n(x_j^2+y_j^2)^{q/2}\Big)^{1/q}=\Big(\sum_{j=1}^n|z_j|^q\Big)^{1/q},$$

where $z_j=x_j+iy_j$, and $B_q^n(\mathbb{C})=\{z\in\mathbb{R}^{2n}:\|z\|_{\ell_q^n(\mathbb{C})}\leqslant 1\}$. Let $H_n=\xi^\perp$ be any complex hyperplane such that $|\xi_1|=\cdots=|\xi_n|$ and $H_1=\eta^\perp$ be such that $\eta_j=0$ for $j\geqslant 2$, where $\xi=(\xi_1,\ldots,\xi_n),\eta=(\eta_1,\ldots,\eta_n)\in\mathbb{C}^n$. In [25], the authors proved that for any $q\in(0,2)$ and any complex hyperplane $H\subseteq\mathbb{C}^n$ the inequalities

$$|B_q^n(\mathbb{C})\cap H_n|\leqslant|B_q^n(\mathbb{C})\cap H|\leqslant|B_q^n(\mathbb{C})\cap H_1| \tag{80}$$

hold true. We will sketch an alternative proof of their result, similar to the proof of Theorem 21. For a complex hyperplane $H=\zeta^\perp$, where $\zeta\in\mathbb{C}^n$, and $\varepsilon>0$ denote by $H_{\mathrm{cyl}}(\varepsilon)$ the cylinder

$$H_{\mathrm{cyl}}(\varepsilon)=\{z\in\mathbb{C}^n:\ |\langle z,\zeta\rangle|<\varepsilon\}. \tag{81}$$

One can prove (see also [40, Corollary 2.5]) that there exists a constant $c(q,n)>0$ such that for every complex hyperplane $H\subseteq\mathbb{C}^n$ we have

$$|B_q^n(\mathbb{C})\cap H|=c(q,n)\lim_{\varepsilon\to0}\frac{1}{\varepsilon^2}\tau_q^n(H_{\mathrm{cyl}}(\varepsilon)), \tag{82}$$

where the measure τ_q^n on $\mathbb{R}^{2n}$ is of the form

$$\mathrm{d}\tau_q^n(x,y)=c_q^n e^{-\sum_{j=1}^n(x_j^2+y_j^2)^{q/2}}\,\mathrm{d}x\,\mathrm{d}y.$$

Writing $e^{-s^{q/2}}=\int_0^\infty e^{-ts}\,\mathrm{d}\mu(t)$ for some measure μ, we deduce that the density of τ_q^n can be written in the form

$$e^{-\sum_{j=1}^n(x_j^2+y_j^2)^{q/2}}=\int_{(0,\infty)^n}e^{-\sum_{j=1}^n t_j(x_j^2+y_j^2)}\,\mathrm{d}\mu_n(t),$$

where $\mu_n = \mu^{\otimes n}$. Therefore, an application of Fubini's theorem and a change of variables imply that there exists a measure ν on $(0,\infty)$ such that for $\nu_n = \nu^{\otimes n}$ and for every Borel set $A \subseteq \mathbb{R}^{2n}$ we can write

$$\tau_q^n(A) = \int_{(0,\infty)^n} \gamma_{2n}(\Delta(y_1, y_1, \ldots, y_n, y_n)^{-1}A)\, \mathrm{d}\nu_n(y), \tag{83}$$

where each coordinate of $y = (y_1, \ldots, y_n)$ is repeated twice. Notice that the image

$$\Delta(y_1, y_1, \ldots, y_n, y_n)^{-1} H_{\mathrm{cyl}}(\varepsilon) = \Big\{ z \in \mathbb{C}^n : \Big| \sum_{j=1}^n \zeta_j y_j z_j \Big| < \varepsilon \Big\}$$

is still a cylinder in $\mathbb{C}^n$ with radius $\big(\sum_{j=1}^n |\zeta_j|^2 y_j^2 \big)^{-1/2} \varepsilon$. Thus, the unitary invariance of complex Gaussian measure and a simple calculation in polar coordinates imply that

$$\gamma_{2n}(\Delta(y_1, y_1, \ldots, y_n, y_n)^{-1} H_{\mathrm{cyl}}(\varepsilon)) = 1 - \exp\Big(-\frac{1}{2} \Big(\sum_{j=1}^n |\zeta_j|^2 y_j^2 \Big)^{-1} \varepsilon^2 \Big). \tag{84}$$

After interchanging limit and integration in (82) and using (83), (84) we deduce that

$$\begin{aligned} |B_q^n(\mathbb{C}) \cap H| &= \frac{c(q,n)}{2} \cdot \int_{(0,\infty)^n} \Big(\sum_{j=1}^n |\zeta_j|^2 y_j^2 \Big)^{-1} \mathrm{d}\nu_n(y) \\ &= \frac{c(q,n)}{2} \cdot \mathbb{E} \Big(\sum_{j=1}^n |\zeta_j|^2 Y_j^2 \Big)^{-1}, \end{aligned}$$

where $Y_1, \ldots, Y_n$ are i.i.d. random variables distributed according to ν. This yields (80) as well as a more general comparison result, similar to Corollary 20, by a direct application of Marshall and Proschan's result (42). $\square$

We note that, in view of Ball's theorem from [6], a Schur monotonicity result for the volume of sections of B_q^n cannot hold in any fixed dimension $n \geqslant 2$ and q large enough. Similarly, according to Szarek's result from [52], the same can be said for the volume of projections of B_q^n for values close to $q = 1$. Finally, we want to stress that a careful look in the previous works [10], [23] and [25] reveals that, even though not stated explicitly, the Schur monotonicity for the volume was established there as well. The new aspect here is the replacement of representations which were Fourier-analytic in flavor by others that exploit the rotational invariance of the Gaussian measure.

Acknowledgements. We would like to thank Apostolos Giannopoulos for providing several useful references and Petros Valettas for his comments regarding Remark 31. We are also indebted to C. Houdré and J. Zinn for communicating the references [4], [32] and [33] to us. Finally, we are very grateful to Assaf Naor for many helpful discussions.

References.

[1] Arora, S. and Kannan, R. (2001). Learning mixtures of arbitrary Gaussians. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing* 247–257. ACM, New York. MR2120323

[2] Artstein-Avidan, S., Giannopoulos, A. and Milman, V. D. (2015). *Asymptotic geometric analysis. Part I. Mathematical Surveys and Monographs* **202**. American Mathematical Society, Providence, RI. MR3331351

[3] Artstein, S., Ball, K. M., Barthe, F. and Naor, A. (2004). Solution of Shannon's problem on the monotonicity of entropy. *J. Amer. Math. Soc.* **17** 975–982 (electronic). MR2083473

[4] Averkamp, R. and Houdré, C. (2003). Wavelet thresholding for non-necessarily Gaussian noise: idealism. *Ann. Statist.* **31** 110–151. MR1962501

[5] Baernstein, A. II and Culverhouse, R. C. (2002). Majorization of sequences, sharp vector Khinchin inequalities, and bisubharmonic functions. *Studia Math.* **152** 231–248. MR1916226

[6] Ball, K. (1986). Cube slicing in $\mathbf{R}^n$. *Proc. Amer. Math. Soc.* **97** 465–473. MR840631

[7] Ball, K., Nayar, P. and Tkocz, T. (2016). A reverse entropy power inequality for log-concave random vectors. *Studia Math.* **235** 17–30. MR3562703

[8] Barthe, F. (1995). Mesures unimodales et sections des boules B_p^n. *C. R. Acad. Sci. Paris Sér. I Math.* **321** 865–868. MR1355843

[9] Barthe, F., Guédon, O., Mendelson, S. and Naor, A. (2005). A probabilistic approach to the geometry of the l_p^n-ball. *Ann. Probab.* **33** 480–513. MR2123199

[10] Barthe, F. and Naor, A. (2002). Hyperplane projections of the unit ball of l_p^n. *Discrete Comput. Geom.* **27** 215–226. MR1880938

[11] Bobkov, S. G. and Chistyakov, G. P. (2015). Entropy power inequality for the Rényi entropy. *IEEE Trans. Inform. Theory* **61** 708–714. MR3332742

[12] Bobkov, S. G. and Houdré, C. (1996). Characterization of Gaussian measures in terms of the isoperimetric property of half-spaces. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **228** 31–38, 356. MR1449845

[13] Bobkov, S. G. and Houdré, C. (1997). Isoperimetric constants for product probability measures. *Ann. Probab.* **25** 184–205. MR1428505

[14] Böröczky, K. J., Lutwak, E., Yang, D. and Zhang, G. (2012). The log-Brunn-Minkowski inequality. *Adv. Math.* **231** 1974–1997. MR2964630

[15] Cordero-Erausquin, D., Fradelizi, M. and Maurey, B. (2004). The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems. *J. Funct. Anal.* **214** 410–427. MR2083308

[16] Dasgupta, S. (1999). Learning mixtures of Gaussians. In *40th Annual Symposium on Foundations of Computer Science (New York, 1999)* 634–644. IEEE Computer Soc., Los Alamitos, CA. MR1917603

[17] Feller, W. (1971). *An introduction to probability theory and its applications. Vol. II. Second edition.* John Wiley & Sons, Inc., New York-London-Sydney. MR0270403

[18] Gozlan, N. and Léonard, C. (2010). Transport inequalities. A survey. *Markov Process. Related Fields* **16** 635–736. MR2895086

[19] Guédon, O., Nayar, P. and Tkocz, T. (2014). Concentration inequalities and geometry of convex bodies. In *Analytical and probabilistic methods in the geometry of convex bodies. IMPAN Lect. Notes* **2** 9–86. Polish Acad. Sci. Inst. Math., Warsaw. MR3329056

[20] Haagerup, U. (1981). The best constants in the Khintchine inequality. *Studia Math.* **70** 231–283 (1982). MR654838

[21] Hardy, G. H., Littlewood, J. E. and Pólya, G. (1988). *Inequalities. Cambridge Mathematical Library*. Cambridge University Press, Cambridge Reprint of the 1952 edition. MR944909

[22] Klartag, B. and Vershynin, R. (2007). Small ball probability and Dvoretzky's theorem. *Israel J. Math.* **157** 193–207. MR2342445

[23] Koldobsky, A. (1998). An application of the Fourier transform to sections of star bodies. *Israel J. Math.* **106** 157–164. MR1656857

[24] Koldobsky, A. (2005). *Fourier analysis in convex geometry. Mathematical Surveys and Monographs* **116**. American Mathematical Society, Providence, RI. MR2132704

[25] Koldobsky, A. and Zymonopoulou, M. (2003). Extremal sections of complex l_p-balls, $0 < p \leqslant 2$. *Studia Math.* **159** 185–194. MR2052217

[26] Koldobsky, A. L. and Montgomery-Smith, S. J. (1996). Inequalities of correlation type for symmetric stable random vectors. *Statist. Probab. Lett.* **28** 91–97. MR1394423

[27] König, H. (2014). On the best constants in the Khintchine inequality for Steinhaus variables. *Israel J. Math.* **203** 23–57. MR3273431

[28] Latała, R. (2002). On some inequalities for Gaussian measures. In *Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002)* 813–822. Higher Ed. Press, Beijing. MR1957087

[29] Latała, R. and Matlak, D. (2017). Royen's proof of the Gaussian correlation inequality. In *Geometric aspects of functional analysis. Lecture Notes in Math.* **2169** 265–275. Springer, Cham. MR3645127

[30] Latała, R. and Oleszkiewicz, K. (1995). A note on sums of independent uniformly distributed random variables. *Colloq. Math.* **68** 197–206. MR1321042

[31] Latała, R. and Oleszkiewicz, K. (2005). Small ball probability estimates in terms of widths. *Studia Math.* **169** 305–314. MR2140804

[32] Lewis, T. M. and Pritchard, G. (1999). Correlation measures. *Electron. Comm. Probab.* **4** 77–85. MR1716783

[33] Lewis, T. M. and Pritchard, G. (2003). Tail properties of correlation measures. *J. Theoret. Probab.* **16** 771–788. MR2009202

[34] Lieb, E. H. (1978). Proof of an entropy conjecture of Wehrl. *Comm. Math. Phys.* **62** 35–41. MR506364

[35] Livne Bar-on, A. (2014). The (B) conjecture for uniform measures in the plane. In *Geometric aspects of functional analysis. Lecture Notes in Math.* **2116** 341–353. Springer, Cham. MR3364696

[36] Marshall, A. W. and Olkin, I. (1979). *Inequalities: theory of majorization and its applications. Mathematics in Science and Engineering* **143**. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London. MR552278

[37] Marshall, A. W. and Proschan, F. (1965). An inequality for convex functions involving majorization. *J. Math. Anal. Appl.* **12** 87–90. MR0185062

[38] Marsiglietti, A. (2016). On the improvement of concavity of convex measures. *Proc. Amer. Math. Soc.* **144** 775–786. MR3430853

[39] Memarian, Y. (2015). On a correlation inequality for Cauchy type measures. *New Zealand J. Math.* **45** 53–64. MR3470260

[40] Meyer, M. and Pajor, A. (1988). Sections of the unit ball of l_p^n. *J. Funct. Anal.* **80** 109–123. MR960226

[41] Nayar, P. and Oleszkiewicz, K. (2012). Khinchine type inequalities with optimal constants via ultra log-concavity. *Positivity* **16** 359–371. MR2929095

[42] Paouris, G. and Valettas, P. (2016). A small deviation inequality for convex functions. Preprint, available at http://arxiv.org/abs/1611.01723. To appear in *Ann. Prob.*.

[43] Paouris, G. and Valettas, P. (2016). Variance estimates and almost Euclidean structure. Preprint, available at https://arxiv.org/abs/1703.10244.

[44] Rachev, S. T. and Rüschendorf, L. (1991). Approximate independence of distributions on spheres and their stability properties. *Ann. Probab.* **19** 1311–1337. MR1112418

[45] Royen, T. (2014). A simple proof of the Gaussian correlation conjecture extended to some multivariate gamma distributions. *Far East J. Theor. Stat.* **48** 139–145. MR3289621

[46] Saroglou, C. (2015). Remarks on the conjectured log-Brunn-Minkowski inequality. *Geom. Dedicata* **177** 353–365. MR3370038

[47] Saroglou, C. (2016). More on logarithmic sums of convex bodies. *Mathematika* **62** 818–841. MR3521355

[48] Schechtman, G. and Zinn, J. (1990). On the volume of the intersection of two L_p^n balls. *Proc. Amer. Math. Soc.* **110** 217–224. MR1015684

[49] Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication.* The University of Illinois Press, Urbana, Ill. MR0032134

[50] Simon, T. (2011). Multiplicative strong unimodality for positive stable laws. *Proc. Amer. Math. Soc.* **139** 2587–2595. MR2784828

[51] Stam, A. J. (1959). Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control* **2** 101–112. MR0109101

[52] Szarek, S. J. (1976). On the best constants in the Khinchin inequality. *Studia Math.* **58** 197–208. MR0430667

[53] Vershynin, R. (2009). Lectures in geometric functional analysis. Lecture notes, available at https://www.math.uci.edu/~rvershyn/papers/GFA-book.pdf.

[54] Weron, A. (1984). Stable processes and measures: a survey. In *Probability theory on vector spaces, III (Lublin, 1983). Lecture Notes in Math.* **1080** 306–364. Springer, Berlin. MR788023

[55] Whittle, P. (1960). Bounds for the moments of linear and quadratic forms in independent variables. *Teor. Verojatnost. i Primenen.* **5** 331–335. MR0133849

[56] Yu, Y. (2008). Letter to the editor: On an inequality of Karlin and Rinott concerning weighted sums of i.i.d. random variables. *Adv. in Appl. Probab.* **40** 1223–1226. MR2488539

Mathematics Department,
Princeton University,
Fine Hall, Washington road, Princeton,
NJ 08544-1000, USA
E-mail: ae3@math.princeton.edu
ttkocz@math.princeton.edu

Wharton Statistics Department,
University of Pennsylvania,
3730 Walnut st., Philadelphia,
PA 19104, USA
E-mail: nayar@mimuw.edu.pl