Entropy jumps for
isotropic log-concave random vectors and spectral gap

by

Keith Ball (Coventry) and Van Hoang Nguyen (Paris)

Abstract. We prove a quantitative dimension-free bound in the Shannon–Stam entropy inequality for the convolution of two log-concave distributions in dimension d in terms of the spectral gap of the density. The method relies on the analysis of the Fisher information production, which is the second derivative of the entropy along the (normalized) heat semigroup. We also discuss consequences of our result in the study of the isotropic constant of log-concave distributions (slicing problem).

1. Introduction. Let X be a random vector in $\mathbb{R}^d$ with density $f :$ $\mathbb{R}^d \to [0, \infty)$, a relation denoted by $X \sim f$. Its entropy is defined to be

$$\text{Ent}(X) = - \int_{\mathbb{R}^d} f \log f$$

provided $\int_{\mathbb{R}^d} f \log_+ f < \infty$. We then say that X *has finite entropy.*

We shall say that a random vector X on $\mathbb{R}^d$, or a probability density f, is *isotropic* if it is centered and has a covariance matrix equal to the identity:

$$\mathbb{E}[X] = \int_{\mathbb{R}^d} x f(x) \, dx = 0$$

and

$$\mathbb{E}[X_i X_j] = \int_{\mathbb{R}^d} x_i x_j f(x) \, dx = \delta_{i,j}, \quad i, j = 1, \dots, d.$$

This normalization can be realized by an affine transformation.

Among random vectors with a given covariance matrix, the corresponding Gaussian has the largest entropy. The gap between the entropy of a random vector X and that of a Gaussian with same covariance matrix is a strong measure of how close X is to being Gaussian. For instance, if X has mean zero and is isotropic with density f, and if G is a standard (nor-

mal) Gaussian vector with density g, then the Pinsker–Csiszár–Kullback inequality (see [20, 11] or [5]) implies that

$$\frac{1}{2}\Big(\int_{\mathbb{R}^d}|f-g|\Big)^2 \le \mathrm{Ent}(G)-\mathrm{Ent}(X).$$

The celebrated Shannon–Stam inequality (see [22, 23]) says that if X and Y are independent identically distributed (i.i.d. for short) random vectors, then the normalized sum $(X+Y)/\sqrt{2}$ has entropy at least as large as that of X and Y:

$$\mathrm{Ent}\Big(\frac{X+Y}{\sqrt{2}}\Big) \ge \mathrm{Ent}(X).$$

Moreover, this inequality is strict if X is not itself a Gaussian random vector. A challenging problem is to quantify this phenomenon, i.e. for fixed covariance matrix (say the identity), bound $\mathrm{Ent}\big(\frac{X+Y}{\sqrt{2}}\big)-\mathrm{Ent}(X)$ from below by a non-negative (and positive outside zero) function of $\mathrm{Ent}(G)-\mathrm{Ent}(X)$. The first result in this direction was obtained by Carlen and Soffer [10] who proved, under technical assumptions, a non-explicit bound based on a compactness argument. Extra assumptions cannot be avoided if one aims at universal entropic estimates: it is easy to construct (by taking a carefully chosen double-bumped Gaussian) a random variable X for which the convolution does not greatly modify the entropy, $\mathrm{Ent}\big(\frac{X+Y}{\sqrt{2}}\big)\simeq\mathrm{Ent}(X)$, but with $\mathrm{Ent}(X)\ll\mathrm{Ent}(G)$.

A surprisingly neat result holds in the case where $X\in\mathbb{R}$ is a random variable with variance 1 and with a density f that satisfies a Poincaré (or spectral gap) inequality in the sense that for some positive c and any smooth function s with $\int_{\mathbb{R}}fs=0$,

$$c\int_{\mathbb{R}}fs^2 \le \int_{\mathbb{R}}f(s')^2.$$

Indeed, we then have

$$(1.1)\qquad \mathrm{Ent}\Big(\frac{X+Y}{\sqrt{2}}\Big)-\mathrm{Ent}(X) \ge \frac{c}{2+2c}(\mathrm{Ent}(G)-\mathrm{Ent}(X))$$

for Y an independent copy of X. This result was proved by Ball, Barthe and Naor in [3] using a variational formula for the Fisher information of a marginal density and spectral analysis to get an information jump in the presence of a spectral gap (see [3, Theorem 2]) and then using a relation between the Fisher information and entropy provided by the adjoint Ornstein–Uhlenbeck semigroup. In the paper [16], Johnson and Barron obtained a result similar to (1.1) under the same hypothesis, but their method is different (at least in details) to that of [3]. In their paper, Johnson and Barron give an upper bound for the L^2 distance of the score function of $(X+Y)/\sqrt{2}$

to the space of additive functions of X and Y, using L^2-orthogonal projections; they eventually use the Poincaré inequality to get the result (see [16, Proposition 2.1 and Lemma 3.1] for details).

The aim of the present paper is to investigate similar results for random vectors, and incidentally to give a new approach to (1.1). A random vector $X \in \mathbb{R}^d$ with density f is said to *satisfy a Poincaré* or *spectral gap inequality* with constant $c > 0$ if for any smooth function s with $\int_{\mathbb{R}^d} fs = 0$,

$$(1.2) \qquad c \int_{\mathbb{R}^d} fs^2 \leq \int_{\mathbb{R}^d} f|\nabla s|^2.$$

The largest constant c in this inequality is indeed the spectral gap for the operator $-L$ on $L^2(f)$ where $Ls := \Delta s - \nabla(\log f) \cdot \nabla s$ for suitable functions; the Poincaré constant $c_p(f)$ refers rather to the inverse of the spectral gap, i.e. to the smallest constant in the inequality

$$(1.3) \qquad \int_{\mathbb{R}^d} fs^2 \leq c_p(f) \int_{\mathbb{R}^d} f|\nabla s|^2.$$

A simplistic adaptation of the argument used in [3] to higher dimensions leads to an inequality of the form (1.1) for random vectors but with an extra dependence on d, the dimension. In the present paper we prove the result without the extra dependence for log-concave random vectors, i.e. those having a density f such that $-\log(f)$ is convex on $\mathbb{R}^n$. It is well known that such random vectors have finite entropy and have a positive spectral gap (see below). This family is central in many high-dimensional problems.

THEOREM 1.1. *Let X be an isotropic log-concave random vector in $\mathbb{R}^d$. Assume its density f satisfies a Poincaré inequality (1.2) with constant $c > 0$. Then, if Y is an independent copy of X, we have*

$$(1.4) \qquad \text{Ent}\left(\frac{X+Y}{\sqrt{2}}\right) - \text{Ent}(X) \geq \frac{c}{4(1+c)}(\text{Ent}(G) - \text{Ent}(X)).$$

Since $c \leq 1$ when f is isotropic, the constant $\frac{c}{4(1+c)}$ may be replaced by $\frac{c}{8}$. The log-concavity assumption will be crucial to the proof of inequality (3.2) below; we do not know whether it holds without this assumption.

As mentioned above, we need to develop a method different from the one of [3]. Our alternative approach relies on the study of *second* derivatives of the entropy along the heat semigroup (or rather along the Ornstein–Uhlenbeck semigroup). Note that it also gives an alternative proof of the one-dimensional case (1.1), up to a numerical (inessential) constant.

Let us make some general comments on log-concave random vectors. If X and Y are i.i.d. random vectors with density f, the normalized sum

$(X + Y)/\sqrt{2}$ has density

$$(1.5) \qquad u \mapsto \int\limits_{\mathbb{R}^d} f\left(\frac{u+v}{\sqrt{2}}\right) f\left(\frac{u-v}{\sqrt{2}}\right) dv,$$

which is a marginal of the joint density on $\mathbb{R}^{2d}$ of the pair (X, Y). It is a consequence of the Brunn–Minkowski inequality (in its functional form due to Prékopa [21]) that log-concave random vectors have log-concave marginals, and hence that if X and Y are log-concave, then so is $(X + Y)/\sqrt{2}$. It is also well known that a log-concave density satisfies a Poincaré inequality for some constant $c > 0$. It was proven by Kannan, Lovász and Simonovitz [17] and independently by Bobkov [6] that if X is an isotropic log-concave random vector, then it satisfies a Poincaré inequality (1.2) with constant C/n for some numerical constant C (thus independent of n and X). Actually, slightly better dependence on n is known. The Kannan–Lovász–Simonovits (KLS) open conjecture states that there exists a universal constant $c > 0$ such that for every n, every isotropic log-concave random vector in $\mathbb{R}^n$ satisfies a Poincaré inequality with constant c. It was noticed some time ago by the first named author, as part of a general program of understanding information theory (and entropy) in the context of convexity in high dimensions, that using (1.4) we can prove that the KLS conjecture implies the celebrated hyperplane (or slicing) conjecture. We shall return to this in the last section.

The organization of the paper is as follows. In the next section, we recall some standard facts about the Ornstein–Uhlenbeck semigroup and about the first derivative of entropy (Fisher information) and the second derivative (information production). Next we investigate how information production behaves under convolution and state a general inequality relating information production of a random vector to the information of a marginal. The subsequent section contains the proof of Theorem 1.1. The final section discusses the connections between entropy jump and the isotropic constant of a log-concave distribution.

2. Classical results on the Ornstein–Uhlenbeck semigroup and Fisher information. For any random vector X with smooth enough density f—we require that $\sqrt{f} \in H^1(\mathbb{R}^n)$, but later the density will even have smoother behavior—its *Fisher information* is defined by

$$J(X) := J(f) := \int\limits_{\mathbb{R}^d} \frac{|\nabla f|^2}{f}.$$

Among random vectors with given covariance matrix, the Gaussian has the smallest Fisher information, as shown by the following straightforward computation: if X is an isotropic mean-zero log-concave random vector with

density f, and G is a standard Gaussian vector with density $g : x \mapsto (1/\sqrt{2\pi})^d e^{-|x|^2/2}$, then

$$J(G) = \int_{\mathbb{R}^d} \frac{|\nabla g|^2}{g} = \frac{1}{\sqrt{2\pi}^d} \int_{\mathbb{R}^d} \sum_{i=1}^d x_i^2 e^{-|x|^2/2} \, dx = d$$

and, by integration by parts,

$$0 \le \int_{\mathbb{R}^d} \left| \frac{\nabla f}{f} + x \right|^2 f(x) \, dx = J(X) - 2 \int_{\mathbb{R}^d} \mathrm{div}(x) f + d = J(X) - J(G).$$

The Fisher information appears as the derivative of the entropy along the Ornstein–Uhlenbeck semigroup, a property central to the works of Bakry and Émery [1] and also to Barron's work [4] on the convergence of entropy in the central limit theorem. The Ornstein–Uhlenbeck semigroup can be constructed in several (equivalent) ways and we choose the following. If X is a random vector with density f and G is a standard Gaussian, independent of X, we consider the random vector $X_t = e^{-t} X + \sqrt{1 - e^{-2t}} \, G$, whose law is the Ornstein–Uhlenbeck evolute at time t of the law of X. More precisely, the density f_t of X_t is the solution of the Fokker–Planck equation with generator L:

$$(2.1) \quad \begin{aligned} & f_0 = f, \\ & \frac{\partial}{\partial t} f_t(x) = L(f_t)(x) := \Delta_x f_t + \mathrm{div}_x(x f_t), \quad \forall t > 0, \forall x \in \mathbb{R}^n. \end{aligned}$$

It is indeed well known that, starting with a (continuous, say) density f, the density f_t is, for $t > 0$, strictly positive, C^∞-smooth on $\mathbb{R}^n$, and that f_t and its derivatives decay exponentially fast to zero at ∞; in particular f_t has a finite Fisher information and it is readily checked that

$$\begin{aligned} \frac{\partial}{\partial t} \mathrm{Ent}(f_t) &= - \int_{\mathbb{R}^d} (L f_t) \log(f_t) \\ &= - \int_{\mathbb{R}^d} f_t \Delta \log f_t - d \int_{\mathbb{R}^d} f_t \\ &= J(f_t) - d. \end{aligned}$$

Hence, we have the classical expression of the entropy gap as the integral of the information gap,

$$(2.2) \quad \mathrm{Ent}(G) - \mathrm{Ent}(X) = \int_0^\infty (J(f_t) - d) \, dt.$$

We refer to Carlen and Soffer [10] for details and precise justifications.

Let us mention for further reference some other nice stability properties of the Ornstein–Uhlenbeck semigroup. It can only improve the spectral gap:

if $X \sim f$ is isotropic and satisfies a Poincaré inequality (1.2) with constant $c > 0$ (actually, $c \in \,]0, 1]$), then f_t satisfies a Poincaré inequality with the same constant $c > 0$. This follows easily from Fubini's theorem, Hölder's inequality and the fact that the Gaussian density has spectral gap of size 1 (see [3]). Next, it is again a consequence of Prékopa's theorem that if X (or f) is log-concave, then so is X_t (or f_t). Finally, it is also classical that the operation of taking marginals commutes with the Ornstein–Uhlenbeck semigroup in the following sense. Let X and Y be two independent random vectors and

$$X_t = e^{-t}X + \sqrt{1 - e^{-2t}}\, G_1 \quad \text{and} \quad Y_t = e^{-t}Y + \sqrt{1 - e^{-2t}}\, G_2$$

their independent evolutes along the Ornstein–Uhlenbeck semigroup, where each G_i $(i = 1, 2)$ is a standard Gaussian vector independent of all the other vectors. Then

$$(2.3) \qquad \frac{X_t + Y_t}{\sqrt{2}} = e^{-t}\frac{X + Y}{\sqrt{2}} + \sqrt{1 - e^{-2t}}\, G$$

where $G = (G_1 + G_2)/\sqrt{2}$ is a standard Gaussian vector.

Throughout the rest of the section, X will be an isotropic log-concave random vector with mean zero, and density f. The density f_t of $X_t = e^{-t}X + \sqrt{1 - e^{-2t}}G$, the evolute of X along the Ornstein–Uhlenbeck semigroup, satisfies (2.1) and takes the form $f_t = e^{-\varphi_t}$ with $\varphi_t := \log(f_t)$ convex on $\mathbb{R}^d$. Its Fisher information will be denoted by

$$(2.4) \qquad \begin{aligned} J(t) := J(f_t) &= \int_{\mathbb{R}^d} \frac{|\nabla f_t|^2}{f_t} = -\int_{\mathbb{R}^d} f_t \Delta \log f_t \\ &= \mathrm{Tr} \int_{\mathbb{R}^d} f_t \, \mathrm{Hess}\, \varphi_t. \end{aligned}$$

We will work with the derivative $\partial_t J(t)$ of the Fisher information along the Ornstein–Uhlenbeck semigroup. The following result is classical in the context of Bakry–Emery's Γ_2 calculus, although it is not usually written in this form which for us will prove useful. We include a proof for completeness.

LEMMA 2.1. *With the previous notation we have*

$$(2.5) \qquad \partial_t J(t) = 2J(t) - 2\,\mathrm{Tr} \int_{\mathbb{R}^d} f_t(\mathrm{Hess}\, \varphi_t)^2.$$

Proof. Denoting by ∂_j the partial derivative (in space) with respect to x_j we have

$$\partial_t J(t) = \sum_{i=1}^{d} \int_{\mathbb{R}^d} 2\frac{\partial_i f_t}{f_t}\partial_i \partial_t f_t - \int_{\mathbb{R}^d} \left(\frac{\partial_i f_t}{f_t}\right)^2 \partial_t f_t$$

$$= \sum_{i=1}^{d} \int_{\mathbb{R}^d} -2\partial_i \varphi_t \partial_i \left[\sum_{j=1}^{d} \partial_j((-\partial_j \varphi_t + x_j)f_t)\right]$$

$$- \int_{\mathbb{R}^d} \sum_{j=1}^{d} \partial_j((-\partial_j \varphi_t + x_j)f_t)(\partial_i \varphi_t)^2.$$

where we took into account that f_t follows (2.1) and $Lg = \sum_j \partial_j(\partial_j g + x_j g)$. Let A and B be the first and the second terms in the above sum, respectively. Then we have, by integration by parts and (2.4),

$$A = 2\sum_{i,j=1}^{d} \int_{\mathbb{R}^d} (\partial_{ij}\varphi_t)\partial_i((-\partial_j \varphi_t + x_j)f_t)$$

$$= 2J(t) - 2\operatorname{Tr}\int_{\mathbb{R}^d} f_t(\operatorname{Hess}\varphi_t)^2 + 2\sum_{i,j=1}^{d}\int_{\mathbb{R}^d} f_t(\partial_{ij}\varphi_t)(\partial_i\varphi_t)(\partial_j\varphi_t - x_j)$$

and

$$B = -\sum_{i=1}^{d}\int_{\mathbb{R}^d}\sum_{j=1}^{d}\partial_j(f_t(-\partial_j\varphi_t + x_j))(\partial_i\varphi_t)^2$$

$$= 2\sum_{i,j=1}^{d}\int_{\mathbb{R}^d} f_t(-\partial_j\varphi_t + x_j)(\partial_i\varphi_t)(\partial_{ij}\varphi_t)$$

$$= -2\sum_{i,j=1}^{d}\int_{\mathbb{R}^d} f_t(\partial_{ij}\varphi_t)(\partial_i\varphi_t)(\partial_j\varphi_t - x_j).$$

Taking the sum of A and B, one gets the result of this lemma. ■

Note that the formula in the previous lemma can be rewritten in the following equivalent form, which is more standard:

$$(2.6) \qquad \partial_t(J(t) - d) = -2(J(t) - d) - 2\operatorname{Tr}\int_{\mathbb{R}^d} f_t\,(\operatorname{Hess}\varphi_t - \operatorname{Id})^2$$

$$\leq -2(J(t) - d).$$

The next lemma will allow us to control the tails of the entropy production.

LEMMA 2.2. *With the previous notation we have*

$$\operatorname{Ent}(G) - \operatorname{Ent}(X) \leq 2\int_0^\infty e^{-2t}(J(t) - d)\,dt.$$

Proof. Integration of inequality (2.6) leads to the following classical Gaussian log-Sobolev inequality:

$$J(t) - d \geq 2(\mathrm{Ent}(G) - \mathrm{Ent}(X_t)), \quad \forall t > 0.$$

By integration by parts, we get

$$\int\limits_0^\infty e^{-2t}(J(t) - d)\,dt \geq \mathrm{Ent}(G) - \mathrm{Ent}(X) - \int\limits_0^\infty e^{-2t}(J(t) - d)\,dt,$$

or equivalently

$$\mathrm{Ent}(G) - \mathrm{Ent}(X) \leq 2\int\limits_0^\infty e^{-2t}(J(t) - d)\,dt. \quad \blacksquare$$

3. A result for the information production of marginals. As we saw in (2.5), the information production $\partial_t J(t)$ along the Ornstein–Uhlenbeck semigroup is given by quantities of the form

$$\mathrm{Tr}\int (\mathrm{Hess}\log f)^2 f.$$

For our argument, we need to analyze how such quantities can be estimated for marginal densities. Assume that Z is a random vector with density $\omega : \mathbb{R}^N \to \mathbb{R}_+$ ($N \geq 1$) and consider the projection $P_E Z$ of Z onto a subspace $E \subset \mathbb{R}^N$. It has a density on $E \simeq \mathbb{R}^{\dim(E)}$ which we denote by h. A useful observation due to Carlen [9] for Fisher information is that

$$J(h) \leq \int \frac{|P_E \nabla \omega|^2}{\omega}.$$

The next result provides an analogue for information production. However, we are able to establish it only in the case of log-concave densities: here is where the restriction in our main theorem comes from. Using it, we can then state the central inequality that will be used in the proof of the main theorem.

LEMMA 3.1. *Let $N \geq 1$ and $\omega = e^{-\phi} : \mathbb{R}^N \to \mathbb{R}_+$ be a smooth positive function. Given a subspace $E \subset \mathbb{R}^N$ define the marginal function on E by*

$$\forall x \in E, \quad h(x) := e^{-\psi(x)} := \int\limits_{E^\perp} \omega(x+y)\,dy = \int\limits_{E^\perp} e^{-\phi(x+y)}\,dy.$$

Denote by P_E the orthogonal projection onto E. Then for every $x \in E$ we have

$$h(x)\,\mathrm{Hess}\,\psi(x) \leq \int\limits_{E^\perp} \omega(x+y)P_E\,\mathrm{Hess}\,\phi(x+y)P_E\,dy$$

in the operator sense (for symmetric operators on E), and if $\mathrm{Hess}\,\psi(x) \geq 0$,

then

$$(3.1) \qquad \mathrm{Tr}[(\mathrm{Hess}\,\psi(x))^2 h(x)] \leq \int_{E^\perp} \mathrm{Tr}[(P_E\,\mathrm{Hess}\,\phi(x+y)P_E)^2]\omega(x+y)\,dy.$$

Therefore, if $\mathrm{Hess}\,\psi \geq 0$ *we have*

$$(3.2) \qquad \int_E \mathrm{Tr}[(\mathrm{Hess}\,\log h)^2]h \leq \int_{\mathbb{R}^N} \mathrm{Tr}[(P_E(\mathrm{Hess}\,\log\omega)P_E)^2]\omega.$$

Proof. We start with the observation that for $x \in E$,

$$(3.3) \qquad \frac{\nabla h(x) \otimes \nabla h(x)}{h(x)} \leq \int_{E^\perp} \frac{P_E\nabla\omega(x+y) \otimes P_E\nabla\omega(x+y)}{\omega(x+y)}\,dy$$

in the symmetric operator sense (on E). Indeed,

$$\nabla h(x) = \int_{E^\perp} P_E\nabla\omega(x+y)\,dy$$

and for any $v \in E$ we have, using the Cauchy–Schwarz inequality,

$$(\nabla h(x), v)^2 \leq \int_{E^\perp} \frac{(\nabla\omega(x+y), v)^2}{\omega(x+y)}\,dy \int_{E^\perp} \omega(x+y)\,dy$$

as claimed. Next, observe that

$$\mathrm{Hess}\,h(x) = \int_{E_\perp} P_E\,\mathrm{Hess}\,\omega(x+y)P_E\,dy$$

and

$$h(x)\,\mathrm{Hess}\,\psi(x) = h(x)\,\mathrm{Hess}(-\log h)(x) = \frac{\nabla h(x) \otimes \nabla h(x)}{h(x)} - \mathrm{Hess}\,h(x).$$

Thus (3.3) leads to the inequality

$$h(x)\,\mathrm{Hess}\,\psi(x)$$
$$\leq \int_{E^\perp} \left(\frac{P_E\nabla\omega(x+y) \otimes P_E\nabla\omega(x+y)}{\omega(x+y)} - P_E\,\mathrm{Hess}\,\omega(x+y)P_E \right) dy$$
$$= \int_{E^\perp} \omega(x+y)\,P_E\,\mathrm{Hess}\,\phi(x+y)P_E\,dy$$

in the operator sense on E, as desired.

Since for symmetric operators $A \geq B \Rightarrow \mathrm{Tr}(AH) \geq \mathrm{Tr}(BH)$ whenever $H \geq 0$, we deduce that, when $\mathrm{Hess}\,\psi(x) \geq 0$,

$$\mathrm{Tr}[(\mathrm{Hess}\,\psi(x))^2 h(x)] \leq \int_{E^\perp} \mathrm{Tr}[P_E\,\mathrm{Hess}\,\phi(x+y)P_E\,\mathrm{Hess}\,\psi(x)]\omega(x+y)\,dy.$$

By the Cauchy–Schwarz inequality (in vector form, for the Hilbert–Schwarz scalar product) we then have

$$\mathrm{Tr}[(\mathrm{Hess}\,\psi(x))^2 h(x)] \leq \sqrt{\int_{E^\perp} \mathrm{Tr}[(P_E \,\mathrm{Hess}\,\phi(x+y)P_E)^2]\omega(x+y)\,dy}$$
$$\times \sqrt{\int_{E^\perp} \mathrm{Tr}[(\mathrm{Hess}\,\psi(x))^2]\omega(x+y)\,dy}.$$

Noting that the second integral equals $\mathrm{Tr}[(\mathrm{Hess}\,\psi(x))^2 h(x)]$, we arrive at inequality (3.1). Integration over E then gives (3.2). ∎

For our argument, we will need the following useful observation.

THEOREM 3.2. *Let X be a log-concave random vector in $\mathbb{R}^d$ with smooth density $f = e^{-\varphi}$ where φ is a convex function on $\mathbb{R}^d$, and let Y be an independent copy of X. Denote by $h = e^{-\psi}$ the density on $\mathbb{R}^d$ of the random vector $(X+Y)/\sqrt{2}$ and put*

$$K = \mathrm{Tr}\Big[\int_{\mathbb{R}^d} (\mathrm{Hess}\,\varphi)^2 f\Big], \quad K_2 = \mathrm{Tr}\Big[\int_{\mathbb{R}^d} (\mathrm{Hess}\,\psi)^2 h\Big]$$

and

$$M = \mathrm{Tr}\Big[\Big(\int_{\mathbb{R}^d} (\mathrm{Hess}\,\varphi)f\Big)^2\Big].$$

Then

$$K_2 \leq \frac{K+M}{2}.$$

Proof. As mentioned earlier, we know by Prékopa's theorem that h is log-concave, i.e. $\mathrm{Hess}\,\psi \geq 0$.

We denote by $\omega(x,y) = f(x)f(y)$ the density of (X,Y) on $\mathbb{R}^d \times \mathbb{R}^d = \mathbb{R}^{2d}$. For $i = 1, \ldots, d$ we set $e_i = (0, \ldots, 1/\sqrt{2}, 0, \ldots, 0, 1/\sqrt{2}, 0, \ldots)$ where the i-th and $(d+i)$-th coordinates are equal to $1/\sqrt{2}$ and the others are zero. Let E be the vector subspace of $\mathbb{R}^{2d}$ spanned by the orthogonal family $\{e_1, \ldots, e_d\}$. We can assume that the density $h = e^{-\psi}$ of the random vector $(X+Y)/\sqrt{2}$ is defined on E by identification of $\mathbb{R}^d$ and E through the orthonormal basis $\{e_i\}$ of E. Then Lemma 3.1 gives

$$K_2 \leq \int_{\mathbb{R}^{2d}} \omega(x,y) \,\mathrm{Tr}\big[[(\mathrm{Hess}(-\log\omega)(x,y)e_i, e_j)]^2_{i,j}\big]\,dx\,dy$$

$$= \int_{\mathbb{R}^{2d}} f(x)f(y) \sum_{i,j=1}^d (\mathrm{Hess}(-\log\omega)(x,y)e_i, e_j)^2\,dx\,dy.$$

Direct computation gives

$$(\mathrm{Hess}(-\log\omega)(x,y)e_i, e_j)^2 = \frac{1}{4}(\partial_{ji}\varphi(x) + \partial_{ji}\varphi(y))^2$$

and hence

$$K_2 \le \sum_{i,j=1}^{d} \frac{1}{4} \int_{\mathbb{R}^{2d}} f(x)f(y)((\partial_{ji}\varphi(x))^2 + 2\partial_{ji}\varphi(x)\partial_{ji}\varphi(y) + (\partial_{ji}\varphi(y))^2)\,dx\,dy$$

$$= \frac{1}{2} \sum_{i,j=1}^{d} \int_{\mathbb{R}^d} f(\partial_{ji}\varphi)^2 + \frac{1}{2} \sum_{i,j=1}^{d} \Big(\int_{\mathbb{R}^d} f\partial_{ji}\varphi \Big)^2 = \frac{1}{2}(K+M). \quad \blacksquare$$

4. Proof of Theorem 1.1. We go back to the situation and the notation of Section 2. That is, X is an isotropic log-concave random vector with mean zero, and density f, and X_t is its evolute along the Ornstein–Uhlenbeck semigroup. The (log-concave) density of X_t is denoted by $f_t = e^{-\varphi_t}$ and we set

$$J(t) := J(X_t) = \text{Tr} \int_{\mathbb{R}^d} f_t \,\text{Hess}\,\varphi_t,$$

$$K(t) := \text{Tr} \int_{\mathbb{R}^d} f_t(\text{Hess}\,\varphi_t)^2 = -\frac{1}{2}e^{2t}\partial_t(e^{-2t}J(t)),$$

where we used Lemma 2.1 for the last equality.

We now consider Z_t, the Ornstein–Uhlenbeck evolute of $(X+Y)/\sqrt{2}$ where Y is an independent copy of X. As mentioned earlier (2.3), $Z_t = (X_t + Y_t)/\sqrt{2}$ where Y_t is an Ornstein–Uhlenbeck evolute of Y independent of X_t. Denote by $h_t = e^{-\psi_t}$ the smooth (log-concave) density of Z_t and set accordingly

$$J_2(t) := J(Z_t) = \text{Tr} \int_{\mathbb{R}^d} h_t \,\text{Hess}\,\psi_t,$$

$$K_2(t) := \text{Tr} \int_{\mathbb{R}^d} h_t(\text{Hess}\,\psi_t)^2 = -\frac{1}{2}e^{2t}\partial_t(e^{-2t}J_2(t)).$$

Theorem 3.2 applied to X_t and $Z_t = (X_t + Y_t)/\sqrt{2}$ then gives

$$K_2(t) \le \frac{K(t)+M(t)}{2} = K(t) - \frac{K(t)-M(t)}{2}$$

where

$$M(t) := \text{Tr}\Big[\Big(\int_{\mathbb{R}^d} (\text{Hess}\,\varphi_t)f_t \Big)^2\Big].$$

This can be rewritten as

$$(4.1) \qquad \partial_t(e^{-2t}(J_2(t) - J(t))) \ge e^{-2t}(K(t) - M(t)).$$

We next claim that

$$(4.2) \qquad K(t) - M(t) \ge \frac{c}{1+c}(K(t) - J(t)).$$

To prove this, recall from Section 2 that f_t satisfies Poincaré's inequality with the same (or better) constant c as f. We apply the Poincaré inequality (1.2) to the density $f_t = e^{-\varphi_t}$ and to the functions

$$s_i(x) = \partial_i \varphi_t(x) - \sum_{j=1}^d x_j \int_{\mathbb{R}^d} (\partial_{ij}\varphi_t) f_t$$

which satisfy $\int s_i f_t = 0$ for $i = 1, \ldots, d$. After summing the inequalities

$$\int_{\mathbb{R}^d} |\nabla s_i|^2 f_t \geq c \int_{\mathbb{R}^d} s_i^2 f_t$$

we find

$$\mathrm{Tr} \int_{\mathbb{R}^d} f_t (\mathrm{Hess}\,\varphi_t)^2 - \mathrm{Tr}\left(\int_{\mathbb{R}^d} f_t \,\mathrm{Hess}\,\varphi_t \right)^2$$

$$\geq c\Big(\mathrm{Tr}\left(\int_{\mathbb{R}^d} f_t \,\mathrm{Hess}\,\varphi_t \right)^2 - \mathrm{Tr} \int_{\mathbb{R}^d} f_t \,\mathrm{Hess}\,\varphi_t \Big).$$

This can be rewritten as $K(t) - M(t) \geq c(M(t) - J(t))$, which is equivalent to the desired inequality (4.2).

Substituting inequality (4.2) in (4.1), we find

$$\partial_t(e^{-2t}(J_2(t) - J(t))) \geq \frac{c}{1+c} e^{-2t}(K(t) - J(t)).$$

Integrating this inequality from t to ∞, we obtain

$$J(t) - J_2(t) \geq \frac{c}{1+c} e^{2t} \int_t^\infty e^{-2s}(K(s) - J(s))\,ds.$$

Hence, using (2.2),

$$\mathrm{Ent}\left(\frac{X+Y}{\sqrt{2}} \right) - \mathrm{Ent}(X) = \int_0^\infty (J(t) - J_2(t))\,dt$$

$$\geq \frac{c}{1+c} \int_0^\infty e^{2t} \int_t^\infty e^{-2s}(K(s) - J(s))\,ds\,dt$$

$$= \frac{c}{2(1+c)} \int_0^\infty (1 - e^{-2t})(K(t) - J(t))\,dt$$

$$= \frac{c}{2(1+c)} \int_0^\infty (1 - e^{-2t})\left(-\frac{1}{2}\partial_t(J(t) - d) \right)dt$$

$$= \frac{c}{2(1+c)} \int_0^\infty e^{-2t}(J(t) - d)\,dt.$$

Applying Lemma 2.2 we get

$$\mathrm{Ent}\left(\frac{X+Y}{\sqrt{2}}\right) - \mathrm{Ent}(X) \geq \frac{c}{4(1+c)}(\mathrm{Ent}(G) - \mathrm{Ent}(X)).$$

This ends the proof of Theorem 1.1.

5. Links with the isotropic constant. The *isotropic constant* of an isotropic log-concave random vector $X \sim f$ on $\mathbb{R}^d$ is defined by

$$L_X := L_f := f(0)^{1/d}.$$

This quantity appears in several high-dimensional problems, and a challenging open problem in asymptotic convex geometry, raised by Bourgain and known as the *slicing* or *hyperplane conjecture*, is whether it is universally bounded (independently of f and d). The best known bound is $L_f \leq cd^{1/4}$ for some universal constant $c > 0$ ([18]). See [2, 19, 18, 12] for background, equivalent formulations and related results.

THEOREM 5.1. *Let X be a random isotropic log-concave vector in $\mathbb{R}^d$. Assume that it satisfies an entropy jump with constant $\kappa \in (0,1)$:*

$$\mathrm{Ent}\left(\frac{X+Y}{\sqrt{2}}\right) - \mathrm{Ent}(X) \geq \kappa(\mathrm{Ent}(G) - \mathrm{Ent}(X))$$

for Y an independent copy of X; by Theorem 1.1, this holds with $\kappa = c/8$ if f satisfies a spectral gap inequality (1.2) with constant $c > 0$. Then

(5.1)
$$L_X \leq e^{2/\kappa}.$$

Note that the bound also reads

$$L_X \leq e^{16\,c_{\mathrm{p}}(X)}$$

in terms of $c_{\mathrm{p}}(X) := c_{\mathrm{p}}(f)$, the Poincaré constant (1.3) of $X \sim f$.

As a consequence, we see that the KLS conjecture (asserting that isotropic log-concave distributions satisfy a Poincaré inequality (1.2) with some universal constant) implies the hyperplane conjecture. In this direction, a better result is known; indeed, Eldan and Klartag [12] recently proved that the *variance conjecture* implies the hyperplane conjecture as well. The variance conjecture asserts that inequality (1.2) for the particular function $s(x) = |x|^2 - \int |y|^2 f(y)\,dy$ holds with a universal constant for every log-concave isotropic distribution f in every dimension d. However, it is worth noting that unlike the Eldan–Klartag result, our estimate above holds at the level of an individual distribution X.

Theorem 5.1 was presented by the first named author in 2003 at a conference in Kiel and then expanded in a series of lectures in 2006 at the conference *Phenomena in High Dimensions* at the I.H.P., as part of a more general program proposing a probabilistic viewpoint on the geometry of

convex bodies in high dimensions. A similar program was also recently and independently proposed by Bobkov and Madiman (see e.g. [7, 8]).

Let us now explain the short and simple argument allowing us to pass from the entropy jump to a bound on the isotropic constant. It relies on a classical rigidity property of isotropic log-concave distributions $X \sim f$ in $\mathbb{R}^d$, namely that up to non-essential linear terms in d, we have $\log f(0) \simeq -\operatorname{Ent}(X)$. The bound

$$(5.2) \qquad -\log f(0) \leq \operatorname{Ent}(X) \leq -\log f(0) + d$$

is for instance implicit in [14] and the easy proof is as follows. Write $f = e^{-\varphi}$ with φ convex. For the lower bound use $\int_{\mathbb{R}^d} xf(x)\,dx = 0$ together with Jensen's inequality to get

$$-\log f(0) = \varphi(0) \leq \int_{\mathbb{R}^d} \varphi(x)f(x)\,dx = \operatorname{Ent}(X).$$

The upper bound combines the convexity of f and an integration by parts as follows:

$$\operatorname{Ent}(X) - \varphi(0) = \int_{\mathbb{R}^d} f(x)[\varphi(x) - \varphi(0)]\,dx \leq \int_{\mathbb{R}^d} f(x)\nabla\varphi(x) \cdot x\,dx = d.$$

Let us mention that in the definition of L_X and in the entropic bounds above, we can replace, up to numerical constants, $f(0)$ by $\|f\|_\infty := \sup_{\mathbb{R}^d} |f|$, since it is known (see [13]) that $\|f\|_\infty \leq e^d f(0)$ for an isotropic log-concave distribution f on $\mathbb{R}^d$.

To finish the proof of (5.1), assume first that $X \sim f$ is symmetric, which means that f is even. If h denotes the density of $(X + Y)/\sqrt{2}$, then $h(x) = 2^{d/2} \int_{\mathbb{R}^d} f(x - y)f(y)\,dy$. It then follows from the log-concavity of f that

$$h(0) = 2^{d/2} \int_{\mathbb{R}^d} f(y)^2\,dy \geq 2^{d/2} \int_{\mathbb{R}^d} f(2y)f(0)\,dy = 2^{-d/2}f(0).$$

Hence, using (5.2) we have

$$(5.3) \qquad \operatorname{Ent}\left(\frac{X + Y}{\sqrt{2}}\right) \leq \frac{d}{2}\log 2 - \log f(0) + d \leq -\log f(0) + \frac{3}{2}d.$$

Let us now go back to the general case where $X \sim f$ is not necessarily symmetric, and consider i.i.d. copies Y, X', Y' of X. Then $(X - X')/\sqrt{2}$ and $(Y - Y')/\sqrt{2}$ are symmetric log-concave isotropic random vectors in $\mathbb{R}^d$, independent and identically distributed according to the density

$$g(x) = 2^{d/2} \int_{\mathbb{R}^d} f(x + y)f(y)\,dy.$$

It follows from the argument above that $g(0) \geq 2^{-d/2}f(0)$. Thus, by the

Shannon–Stam inequality

$$\frac{1}{2}\operatorname{Ent}(Z) + \frac{1}{2}\operatorname{Ent}(U) \le \operatorname{Ent}\left(\frac{Z+U}{\sqrt{2}}\right)$$

for $Z = (X+Y)/\sqrt{2}$ and $U = -(X'+Y')/\sqrt{2}$ two independent random vectors, and by the bound (5.3) obtained in the symmetric case, one gets

$$\operatorname{Ent}\left(\frac{X+Y}{\sqrt{2}}\right) \le \operatorname{Ent}\left(\frac{X+Y-X'-Y'}{\sqrt{4}}\right) = \operatorname{Ent}\left(\frac{\frac{X-X'}{\sqrt{2}} + \frac{Y-Y'}{\sqrt{2}}}{\sqrt{2}}\right)$$

$$\le -\log g(0) + \frac{3}{2}d \le -\log f(0) + 2d.$$

On the other hand, the assumption on the entropy jump implies

$$(1-\kappa)\operatorname{Ent}(X) \le \operatorname{Ent}\left(\frac{X+Y}{\sqrt{2}}\right) - \kappa\operatorname{Ent}(G) \le \operatorname{Ent}\left(\frac{X+Y}{\sqrt{2}}\right)$$

since $\operatorname{Ent}(G) = (d/2)\log 2\pi e \ge 0$. Therefore, using again (5.2) we get

$$(1-\kappa)[-\log f(0)] \le -\log f(0) + 2d.$$

This implies

$$\kappa \log f(0) \le 2d$$

and hence the desired bound (5.1).

References

[1] D. Bakry and M. Émery, *Diffusions hypercontractives*, in: Séminaire de probabilités, XIX, 1984/84, J. Azéma and M. Yor (eds.), Lecture Notes in Math. 1123, Springer, Berlin, 1985, 177–206.

[2] K. Ball, *Logarithmically concave functions and sections of convex sets in* $\mathbf{R}^n$, Studia Math. 88 (1988), 69–84.

[3] K. Ball, F. Barthe and A. Naor, *Entropy jumps in the presence of a spectral gap*, Duke Math. J. 119 (2003), 41–63.

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

[5] N. M. Blachman, *The convolution inequality for entropy powers*, IEEE Trans. Inform. Theory 11 (1965), 267–271.

[6] S. G. Bobkov, *Isoperimetric and analytic inequalities for log-concave probability measures*, Ann. Probab. 27 (1999), 1903–1921.

[7] S. G. Bobkov and M. Madiman, *Dimensional behaviour of entropy and information*, C. R. Math. Acad. Sci. Paris 349 (2011), 201–204.

[8] S. G. Bobkov and M. Madiman, *The entropy per coordinate of a random vector is highly constrained under convexity conditions*, IEEE Trans. Inform. Theory 57 (2011), 4940–4954.

[9]   E. A. Carlen, *Superadditivity of Fisher's information and logarithmic Sobolev in-equalities*, J. Funct. Anal. 101 (1991), 194–211.
[10]   E. A. Carlen and A. Soffer, *Entropy production by block variable summation and central limit theorems*, Comm. Math. Phys. 140 (1991), 339–371.
[11]   I. Csiszár, *Informationstheoretische Konvergenzbegriffe im Raum der Wahrschein-lichkeitsverteilungen*, Magyar Tud. Akad. Mat. Kutató Int. Kozl. 7 (1962), 137–158.
[12]   R. Eldan and B. Klartag, *Approximately Gaussian Marginals and the Hyperplane Conjecture*, in: Concentration, Functional Inequalities and Isoperimetry, Contemp. Math. 545, Amer. Math. Soc., Providence, RI, 2011, 55–68.
[13]   M. Fradelizi, *Sections of convex bodies through their centroid*, Arch. Math. (Basel) 69 (1997), 515–522.
[14]   M. Fradelizi and M. Meyer, *Increasing functions and inverse Santaló inequality for unconditional functions*, Positivity 12 (2008), 407–420.
[15]   L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. 97 (1975), 1061–1083.
[16]   O. Johnson and A. R. Barron, *Fisher information inequalities and the central limit theorem*, Probab. Theory Related Fields 129 (2004), 391–409.
[17]   R. Kannan, L. Lovász, and M. Simonovits, *Isoperimetric problems for convex bodies and a localization lemma*, Discrete Comput. Geom. 13 (1995), 541–559.
[18]   B. Klartag, *On convex perturbations with a bounded isotropic constant*, Geom. Funct. Anal. 16 (2006), 1274–1290.
[19]   V. D. Milman and A. Pajor, *Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space*, in: Geometric Aspects of Functional Analysis (1987–88), Lecture Notes in Math. 1376, Springer, Berlin, 1989, 64–104.
[20]   M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden-Day, San Francisco, 1964.
[21]   A. Prékopa, *On logarithmic concave measures and functions*, Acta Sci. Math. (Szeged) 34 (1973), 335–343.
[22]   C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, Univ. of Illinois Press, Urbana, IL, 1949.
[23]   A. J. Stam, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Inform. and Control 2 (1959), 101–112.

Keith Ball   Van Hoang Nguyen
Institute of Mathematics   Institut de Mathématiques de Jussieu
University of Warwick   UPMC
Coventry, CV4 7AL, UK   4 place Jussieu
E-mail: kmb120205@googlemail.com   75252 Paris, France
                                        E-mail: vanhoang@math.jussieu.fr