Stability and Equality in the Entropy Power Inequality and in one of its Discrete Counterparts

Lampros Gavalakis and Ioannis Kontoyiannis
University of Cambridge
{lg560, ik355}@cam.ac.uk

***Abstract*—We show that there is equality in Shannon's Entropy Power Inequality (EPI) if and only if the random variables involved are Gaussian, assuming nothing beyond the existence of differential entropies. This is done by justifying de Bruijn's identity without a second moment assumption. Part of the proof also relies on a re-examination of an example of Bobkov and Chistyakov (2015), which shows that there exists a random variable X with finite differential entropy $h(X)$, such that $h(X+Y)=\infty$ for any independent random variable Y with finite entropy. We prove that either X has this property, or $h(X+Y)$ is finite for any independent Y that does not have this property. Using this, we prove the continuity of $t \mapsto h(X+\sqrt{t}Z)$ at $t=0$, where $Z \sim \mathcal{N}(0,1)$ is independent of X, under minimal assumptions. We then establish two stability results: A qualitative stability result for Shannon's EPI in terms of weak convergence under very mild moment conditions, and a quantitative stability result in Tao's discrete analogue of the EPI under log-concavity. The proof for the first stability result is based on a compactness argument, while the proof of the second uses the Cheeger inequality and leverages concentration properties of discrete log-concave distributions.**

All of our results are stated without their proofs; these can be found in the full version of the paper at [19].

I. Shannon's Entropy Power Inequality

One of the equivalent formulations of Shannon's celebrated entropy power inequality (EPI) states that, for any pair of independent random variables X, Y in $\mathbb{R}$ and any $\lambda \in (0,1)$,

$$h(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \geq \lambda h(X) + (1-\lambda)h(Y) \tag{1}$$

where $h(X)$ denotes the (differential) entropy of a random variable X having density f,

$$h(X) = h(f) = -\int_{\mathbb{R}} f(x)\log f(x)dx, \tag{2}$$

with the convention that we set $h(X) = -\infty$ whenever the integral in (2) does not exist, including in the case when X does not have a density with respect to Lebesgue measure on $\mathbb{R}$. In particular, when we say that the differential entropy $h(X)$ of some random variable X is finite, we mean that the integral exists and $-\infty < h(X) < \infty$, while "$h(X) < \infty$" also includes the cases when the integral does not exist and when it does exist and equals $-\infty$. [Logarithms are natural logarithms throughout.]

The first proof of the EPI, after it was stated in Shannon's paper [42], was given by Stam [43] in 1959 and made use of de Bruijn's identity,

$$\frac{d}{dt}h(X+\sqrt{t}Z) = \frac{1}{2}I(X+\sqrt{t}Z), \quad t>0. \tag{3}$$

Here, Z is a standard Gaussian independent of X, and $I(X)$ denotes the Fisher information of X, where

$$I(X) = \int_{\mathbb{R}} \frac{\big(f'(x)\big)^2}{f(x)}dx, \tag{4}$$

whenever X has an absolutely continuous density f; otherwise, we set $I(X)=\infty$, c.f. [8].

Stam's proof was revisited by Blachman [7] and it was also presented in the review of Dembo, Cover and Thomas [14]. Using de Bruijn's identity (3), the EPI is derived as a consequence of the convolution inequality for Fisher information, also known as Stam's inequality [7], [43]:

$$I(X+Y)^{-1} \geq I(X)^{-1} + I(Y)^{-1},$$

where X, Y are arbitrary independent random variables. The proof of de Bruijn's identity (3) requires justification of exchange of differentiation and expectation. As has been pointed out before [30], [41], this was only done in 1984 by Barron [6], under the assumption that X has finite variance. Around the same time, the Bakry-Eméry theory was developed [2], where a related representation for the time derivative of the relative entropy $D(P_t\|Q)$ between the time-t distribution P_t of a diffusion and its invariant measure Q was obtained. A proof of the EPI using de Burijn's identity also appeared in a slightly more general form in Carlen and Soffer [10], where a qualitative stability result is obtained as well. We will discuss this in more detail in the following section. More recently, another proof of the EPI in the same spirit, using a closely related derivative identity for the mutual information, was given by Verdú and Guo [46], again assuming finite variance.

One of the main contributions of the present work is the justification of de Bruijn's identity (3) under minimal assumptions and in particular under no moment conditions, see Theorem 4.

Another part of Stam's proof which also uses the finite variance assumption, is the continuity of $t \mapsto h(X+\sqrt{t}Z)$ at $t=0$. We will show in Theorem 3 that this can be justified without moment conditions but, as we will see, this continuity fails without mild additional assumptions.

A different classical proof of the EPI is due to Lieb [34] (see also [14]), using the sharp form of Young's inequality, which yields a family of inequalities for Rényi entropies. However, this proof requires integrability of some power of the densities, i.e., $\int f^\alpha < \infty$ for some $\alpha > 1$, in order to obtain the EPI for differential entropy in the limit as $\alpha \to 1$. One can remove this condition to obtain the EPI for any pair of random variables using a truncation argument as in Bobkov and Chistyakov [9]. But because of this truncation, this method of proof does not settle the equality case without the above integrability assumption.

A number of other proofs of the EPI have appeared in the literature, under a variety of different assumptions. A simple proof, which also generalizes to give entropy monotonicity along the central limit theorem (CLT), is due to Madiman and Barron [35], under the finite variance assumption. Several different proofs were proposed by Rioul [38]–[41]. In particular, the equality case is settled in [41], assuming (among other conditions) that the densities of X and Y are everywhere positive and continuous.

Szarek and Voiculescu [44] gave a nice, geometric and intuitive proof using the Brunn-Minkowski inequality and a rearrangement inequality by Brascamp, Lieb and Luttinger. Due to the limiting nature of this argument, the equality case is not settled there either. Anantharam, Jog and Nair [1] obtained a proof that unifies the EPI and the Brascamp-Lieb inequality using a doubling trick, assuming finite second moments. Courtade [12] also provided a simple proof for the monotonicity of entropy along the CLT using the maximal correlation, assuming finite variances and smooth densities. Finally, a recent, remarkable proof using the Föllmer process is due to Lehec [32], which assumes finite second moments. A similar idea was used by Eldan and Mikulincer [16] to obtain stability under log-concavity. We will revisit this topic in the following section.

Since the proofs in the literature only settle the equality case of the EPI in restricted cases (assuming either finite second moments or under regularity assumptions on the densities of X and Y), it is natural to ask: *Assuming only finiteness of $h(X)$ and $h(Y)$, does equality in the EPI* (1) *hold if and only if X and Y are Gaussian?* We answer this question in the affirmative, using our general version of de Bruijn's identity, a result which may also be of independent interest.

Before stating this in Theorem 5, we note that, as Bobkov and Chistyakov [9] point out, one needs to be careful even when formulating the EPI (1): there exist random variables X, Y such that the integrals in the definitions of $h(X)$ and $h(Y)$ exist, while $h(X+Y)$ does not. Under the convention that $h(X) = -\infty$ when the integral does not exist, (1) holds true for any pair of independent random variables X, Y.

Furthermore, Bobkov and Chistyakov showed that the following "discontinuity" phenomenon may occur [9, Proposition V.8]:

Proposition 1 ([9])**.** *There exists a random variable X such that $-\infty < h(X) < \infty$, while $h(X+Y) = \infty$ for every independent random variable Y with $h(Y) > -\infty$.*

This example shows that $t \mapsto h(X + \sqrt{t}Z)$, where Z is a standard Gaussian independent of X, need not be continuous at $t = 0$. Establishing continuity at $t = 0$ is a key step in Stam's proof of the EPI, and it fails when no conditions other than the existence of $h(X), h(Y)$ are imposed. For the same reason, de Bruijn's identity (3) may also fail, as its right-hand side is finite.

Our first main contribution, Theorem 2 below, shows that the family of counterexamples provided by Bobkov and Chistyakov above is the "only" pathological case. For convenience, we first define the class of random variables having the property of the counterexample of Proposition 1:

$$\mathcal{C}_{\mathrm{BC}} := \{X : h(X) \text{ is finite and } h(X+Y) = \infty \\ \text{for any independent } Y \text{ with } h(Y) > -\infty\}.$$

Theorem 2. *Let X be a random variable in $\mathbb{R}$ whose differential entropy exists and is finite. There are only two possibilities:*

1) *Either $X \in \mathcal{C}_{\mathrm{BC}}$, or*
2) *$h(X+Y) < \infty$ for any independent random variable $Y \notin \mathcal{C}_{\mathrm{BC}}$.*

In addition, we show that $X \notin \mathcal{C}_{\mathrm{BC}}$ is enough to justify the continuity of entropy under Gaussian perturbation:

Theorem 3 (Entropy continuity under Gaussian perturbation)**.** *Let X be a random variable in $\mathbb{R}$ with finite differential entropy, for which there exists an independent random variable Y with finite differential entropy such that $h(X+Y) < \infty$. Then, if $Z \sim \mathcal{N}(0,1)$ is independent of X,*

$$\lim_{t \downarrow 0} h(X + \sqrt{t}Z) = h(X).$$

The proofs exploit a submodularity-for-sums inequality for entropy [31] and use a truncation argument, together with some uniform estimates on the mutual information, and a few careful applications of dominated convergence.

Our next goal is to derive de Bruijn's identity (3) under the weakest possible conditions. As mentioned, we see from Proposition 1 that the existence and finiteness of $h(X)$ is not enough. Nevertheless, we show:

Theorem 4 (De Bruijn's identity without finite variance)**.** *Let X be a random variable in $\mathbb{R}$ with $h(X) < \infty$, for which there exists an independent random variable Y with finite differential entropy such that $h(X+Y)$ exists and is finite. Then, if $Z \sim \mathcal{N}(0,1)$ is independent of X,*

$$\frac{d}{dt} h(X + \sqrt{t}Z) = \frac{1}{2} I(X + \sqrt{t}Z), \quad t > 0.$$

The key idea in the proof of Theorem 4 is to replace the step in the proof of Barron [6] where the finite variance assumption is used, with the finiteness of $h(X + \sqrt{t_0}Z)$ for some $t_0 > 0$, which is a consequence of the submodularity-for-sums inequality for entropy. A data processing argument together with an application of dominated convergence complete the proof.

Thus, we are able to prove the following general version of the EPI. We require no assumptions for the EPI to hold and we settle the equality case under the minimal assumption that the entropies exist and are finite. Clearly, if both sides are infinite, equality can be achieved trivially without the random variables being Gaussian.

Theorem 5 (EPI)**.** *Let X, Y be independent random variables on $\mathbb{R}$. Then, for any $\lambda \in (0,1)$,*

$$h(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \geq \lambda h(X) + (1-\lambda)h(Y). \quad (5)$$

If $h(X)$ and $h(Y)$ exist and are finite, there is equality if and only if X and Y are Gaussian with the same (finite) variance.

To the best of our knowledge, this is the first characterization of the equality case in the EPI under no assumption other than existence of entropies.

As noted above, a version of de Bruijn's identity for the derivative of relative entropy can be established using the Bakry-Émery theory [3, Proposition 5.2.2]. For a diffusion $\{X_t\}$ on $\mathbb{R}$, this is proved for every function f in the associated Dirichlet domain $\mathcal{D}(\mathcal{E})$. The present setting corresponds to the Ornstein-Uhlenbeck process on the real line, with $f = \frac{d\mu}{d\gamma}$, where μ denotes the law of X and γ is a Gaussian measure. However, since we are not assuming finite variance for X, it is not clear how to translate [3, Proposition 5.2.2] to the derivative of the entropy itself. Moreover, our assumptions are strictly weaker: here, the Dirichlet domain reduces to the set of functions having a weak derivative in $L^2(\gamma)$ [3, p. 126], which excludes, for example, functions with jumps. Our result does not require any such regularity assumptions and includes, for example, the case where X has a piecewise constant density.

Finally, we state two simple consequences of Theorem 4. Sometimes, in applications arising in the study of generative models in machine learning, it is of interest to differentiate the entropy of a diffusion with discrete initial distribution. Corollary 6 says that de Bruijn's identity is valid for any discrete X with finite discrete entropy $H(X)$. Recall that the discrete (Shannon) entropy of a discrete random variable X with probability mass function p on a finite or countably infinite set A is

$$H(X) = H(p) = -\sum_{a \in A} p(a) \log p(a). \quad (6)$$

Corollary 6. *Let X be a discrete real-valued random variable with finite entropy $H(X)$. Then, if $Z \sim \mathcal{N}(0,1)$ is independent of X,*

$$\frac{d}{dt} h(X + \sqrt{t}Z) = \frac{1}{2} I(X + \sqrt{t}Z), \quad t > 0.$$

A combination of Theorems 3 and 4 with the mean value theorem gives de Bruijn's identity at $t = 0$. Recall that, by convention, we set the Fisher information $I(X) = \infty$ for any random variable that does not have an absolutely continuous density.

Corollary 7. *Let X be a random variable in $\mathbb{R}$ with $h(X) < \infty$, for which there exists an independent random variable Y with finite differential entropy such that $h(X+Y)$ exists and is finite. Then, if $Z \sim \mathcal{N}(0,1)$ is independent of X,*

$$\frac{d}{dt} h(X + \sqrt{t}Z)\Big|_{t=0^+} = \frac{1}{2} I(X). \quad (7)$$

II. (In)stability of Shannon's EPI

The next question we address is that of stability in Shannon's EPI: if the deficit in the EPI (1) is small, are the random variables X, Y close (in some appropriate sense) to the extremizers, i.e., to Gaussians? To make this precise, for $\lambda \in (0,1)$ let

$$\delta_{\mathrm{EPI},\lambda}(X,Y) := h(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) - \lambda h(X) - (1-\lambda)h(Y).$$

Does

$$\delta_{\mathrm{EPI},\lambda}(X,Y) \leq \delta,$$

imply that

$$d(X, G_X), d(Y, \mathrm{G}_Y) \leq \epsilon(\delta),$$

for some Gaussians G_X, G_Y, some "distance" measure d, and with $\epsilon(\delta) \to 0$ as $\delta \to 0$? Is it possible to obtain quantitative estimates for $\epsilon(\delta)$?

The first (positive) stability result was obtained by Carlen and Soffer in [10], where it was shown that, among isotropic random vectors that satisfy a uniform bound on the Fisher information and a uniform bound on the decay of the second moment tails, $\delta_{\mathrm{EPI},\lambda} \to 0$ implies convergence to Gaussianity in relative entropy. In the reverse direction, Courtade, Fathi and Pananjady [13] constructed a sequence of random variables that satisfy $\delta_{\mathrm{EPI},\lambda} \to 0$, while their relative entropy from any Gaussian is bounded away from zero. These random variables do not satisfy Carlen and Soffer's uniform bound on the second moment tails, but they have unit variance and bounded Fisher information. A quantitative stability result for uniformly log-concave random variables is also established in [13].

Ball, Barthe and Naor [4] proved that, if X has finite Poincaré constant $C_P(X)$, then

$$D(X) \leq \frac{2C_P(X) + 2\sigma^2}{\sigma^2} \delta_{\mathrm{EPI},\frac{1}{2}}(X,X), \quad (8)$$

where $\delta_{\mathrm{EPI},\frac{1}{2}}(X,X)$ refers to the case when X and Y have the same distribution, σ^2 is the variance of X, and $D(X)$ denotes the relative entropy between X and a Gaussian with the same mean and variance as X. This was generalized to higher dimensions under the extra assumption that X is log-concave by Ball and Nguyen [5]. Related bounds were also established in [31], [36]; in particular, using the results of Johnson and Barron [25], it was shown in [31] that, if X has finite Poincaré constant $C_P(X)$, then

$$D(X) \leq \frac{2C_P(X) + \sigma^2}{\sigma^2} \delta_{\mathrm{EPI},\frac{1}{2}}(X,X). \quad (9)$$

Further discussion of the relation between $\delta_{\mathrm{EPI},\frac{1}{2}}(X,X)$ and the closely related "entropic doubling constant" is given in [21], [33].

Eldan and Mikulincer [16] used an adaptation of Lehec's idea [32] to generalize (8) for any λ (instead of only considering $\lambda = \frac{1}{2}$) and to obtain improved constants in the case where the random variables are already close to Gaussian. This result played a key role in the recent breakthrough resolution of Bourgain's slicing problem [24], [29].

Recently, a weak qualitative stability result was established in [21], in a similar spirit as that in [10], but without the Fisher information assumption, and in terms of weak convergence rather than convergence in relative entropy.

Although the counterexample of [13] says that stability may fail for strong distances such as relative entropy, it is natural to ask whether we may still have stability in the sense of weak convergence. The result of [21] does not completely answer this, as the uniform integrability assumption in that result, which is necessary in the proof due to the use of the result of [10], is not satisfied by the counterexample of [13].

Our main stability result is only qualitative and in the weakest possible distance, the Lévy metric d_L [15], which metrizes weak convergence, but it only assumes finite kth moments for some $k > 1$:

Theorem 8. *Suppose X, Y are random variables with finite differential entropies, $-\infty < h(X), h(Y) < \infty$, and*

$$\mathbb{E}|X|^k, \mathbb{E}|Y|^k < \infty \text{ for some } k > 1.$$

Fix $\lambda \in (0,1)$. Then

$$\begin{aligned} &\text{for each } \epsilon > 0 \text{ there is a } \delta > 0 \text{ s.t.} \\ &\delta_{\mathrm{EPI},\lambda}(X,Y) < \delta \;\text{ implies }\; d_{\mathrm{L}}(X,G_1), d_{\mathrm{L}}(Y,G_2) < \epsilon, \end{aligned} \tag{10}$$

for some Gaussian random variables G_1, G_2 with the same variance.

Equivalently, suppose $\lambda \in (0,1)$ and $\{(X_n, Y_n)\}$ is a sequence of pairs of random variables such that, for each n, X_n, Y_n are independent with finite differential entropies. If $\{X_n\}$ and $\{Y_n\}$ have uniformly bounded kth absolute moments for some $k > 1$ and $\delta_{\mathrm{EPI},\lambda}(X_n, Y_n) \to 0$ as $n \to \infty$, then the limits of any subsequence along which $\{X_n\}$ and $\{Y_n\}$ both converge weakly, are Gaussian with the same variance.

Note that the conclusion of Theorem 8 does not imply that the random variables have a weak limit if the deficit in the EPI vanishes; in fact this may well not hold. It is also straightforward to check that the counterexample of [13] satisfies the assumption of our Theorem 8, since the densities in that example all have unit variance. The proof is based on a compactness argument. After smoothing, the moment condition ensures convergence of differential entropies and we obtain the result by the equality case in the EPI via Theorem 5.

III. Stability of Tao's discrete EPI

Obtaining discrete analogues of the EPI is a subtle task and it is not ever clear what is the most natural discrete version of Shannon's EPI. Tao [45] proved the following discrete analogue for independent and identically distributed (i.i.d.) discrete random variables taking values in a torsion-free group:

Theorem 9 ([45])**.** *Let X_1, X_2 be i.i.d. random variables with values in a discrete subset of a torsion-free group G. Then*

$$H(X_1 + X_2) \geq H(X_1) + \frac{1}{2}\log 2 - o(1),$$

as $H(X_1) \to \infty$.

Here H denotes the discrete (Shannon) entropy, recall (6). This can be seen as a discrete analogue of (1) for $\lambda = \frac{1}{2}$, which by scaling, is the same as

$$h(X_1 + X_2) \geq h(X_1) + \frac{1}{2}\log 2.$$

It is easy to see that without the $o(1)$ term, the inequality

$$H(X_1 + X_2) \geq H(X_1) + \frac{1}{2}\log 2,$$

can fail if $H(X_1)$ is small enough.

In the case of integer-valued random variables, finite bounds for the $o(1)$ term were established in [17], [18], along with generalizations to sums of more than two random variables, under the discrete log-concavity assumption:

Theorem 10 ([18, Theorem 2, case $n = 1$])**.** *Let X_1, X_2 be i.i.d. log-concave random variables in $\mathbb{Z}$. There is an absolute constant C such that,*

$$H(X_1 + X_2) \geq H(X_1) + \frac{1}{2}\log 2 - C\frac{\log\sigma}{\sigma}, \tag{11}$$

provided that $\sigma^2 := \mathrm{Var}(X_1)$ is large enough.

Our main result in the discrete setting provides a quantitative stability estimate of (11):

Theorem 11. *Suppose X has a log-concave distribution on the integers, variance $\mathrm{Var}(X) = \sigma^2$ and X_1, X_2 are i.i.d. copies of X. Denote by $Z^{(\mathbb{Z})}$ a Gaussian with the same mean and variance as X, discretized on $\mathbb{Z}$. Assume $\sigma \geq 1547$. Then there are absolute constants C_1, C_2 such that*

$$\begin{aligned} &D(X\|Z^{(\mathbb{Z})}) \\ &\leq C_1\Big(H(X_1+X_2) - H(X_1) - \frac{1}{2}\log 2\Big) + C_2\frac{\log\sigma}{\sigma}, \end{aligned} \tag{12}$$

where $D(\cdot\|\cdot)$ denotes the relative entropy.

From the proof of Theorem 11 it can be seen that we may take $C_1 = 876489$. An explicit value for C_2 may also be extracted, but it will be at least 10^{10}, so we do not pursue this here. We expect that these constants can be improved significantly.

By a discretized Gaussian $Z^{(\mathbb{Z})}$ in Theorem 11 (also referred to as quantized Gaussian in [20]) we mean

$$\mathbb{P}(Z^{(\mathbb{Z})} = k) = \int_{[k,k+1)} \varphi_{\mu,\sigma^2}(x)dx, \quad k \in \mathbb{Z},$$

where φ_{μ,σ^2} denotes the Gaussian density with the same mean μ and variance σ^2 as X.

Theorem 11 says that if the discrete EPI (11) is close to equality, and the if variance of X_1 is large enough (a condition which is, in general, weaker than the entropy being large), then X_1 is close in relative entropy to a Gaussian discretized on the integers. It can also be viewed as a discrete analogue of the continuous stability results of Ball, Barthe and Naor (8) and of Kontoyiannis-Madiman (9).

Our strategy for proving Theorem 11 is to use (9) applied to $X+U$, where $U \sim \mathcal{U}[0,1]$, and to approximate the continuous EPI deficit by the discrete one using the results of [18]. Note that, since the density of $X+U$ is piecewise constant, it is not log-concave. Nevertheless, we show that it has a finite Poincaré constant (see, e.g., [27] for definitions) bounded by an absolute constant times the variance of X:

Proposition 12. *Let X be a discrete log-concave random variable on the integers and let U be a continuous uniform random variable on the unit interval, independent of X. Assume $\sigma^2 := \mathrm{Var}(X) \geq 4$. Then*

$$C_P(X+U) \leq 438244 \cdot \sigma^2,$$

where $C_P(\cdot)$ denotes the Poincaré constant.

This bound is analogous to the continuous case, up to the large absolute constant. In the proof we use the Cheeger inequality [11],

$$C_P(\nu) \leq \frac{4}{\mathrm{Is}^2(\nu)},$$

where $\mathrm{Is}(\nu)$ denotes the isoperimetric constant of a probability measure ν. A useful identity obtained by Bobkov and Houdré in the one-dimensional case then allows us to exploit the concentration properties of discrete log-concave measures to bound the isoperimetric constant.

We mention in passing that bounding the Poincaré constant (or the inverse of the isoperimetric constant) from above in high dimensions is related to the Kannan-Lovász-Simonovits conjecture [26], which states that the Poincaré constant of any (continuous) log-concave random vector is bounded above by an absolute constant, independent of the dimension. The best known upper bound to date is $C\sqrt{\log n}$ due to Klartag [28]. Of course, our Proposition 12 is only in dimension 1.

Another motivation for considering estimates as in (12) comes from the recent proof of Marton's conjecture [22], [23]. The equivalent entropic formulation of the polynomial Freiman-Ruzsa (PFR) conjecture [22, Theorem 1.8] can be seen as a stability estimate in the lower bound

$$H(X_1+X_2) - \frac{1}{2}H(X_1) - \frac{1}{2}H(X_2) \geq 0$$

when X_1, X_2 take values in $\mathbb{F}_2^n$, in terms of a "distance" known as Ruzsa distance. The extremizers in this case are uniform distributions on finite subgroups. Obtaining analogous estimates in torsion-free groups is related to PFR over $\mathbb{Z}$, which remains open. An improvement of the best known bounds in this case is given in the recent preprint [37], which also discusses the potential usefulness of discrete stability estimates in Ruzsa distance.

Acknowledgment

This work was supported in part by the EPSRC-funded INFORMED-AI project EP/Y028732/1.

References

[1] V. Anantharam, V. Jog, and C. Nair. Unifying the Brascamp-Lieb inequality and the entropy power inequality. *IEEE Trans. Inform. Theory*, 68(12):7665–7684, December 2022.

[2] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84*, pages 177–206. Springer, Berlin, 1985.

[3] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften*. Springer, Cham, Germany, 2014.

[4] K. Ball, F. Barthe, and A. Naor. Entropy jumps in the presence of a spectral gap. *Duke Math. J.*, 119(1):41–63, July 2003.

[5] K. Ball and V.H. Nguyen. Entropy jumps for isotropic log-concave random vectors and spectral gap. *Studia Mathematica*, 213(1):81–96, 2012.

[6] A.R. Barron. Monotonic central limit theorem for densities. NSF Technical Report no. 50, Department of Statistics, Stanford University, March 1984. Available at `statistics.stanford.edu/resources/technical-reports`.

[7] N.M. Blachman. The convolution inequality for entropy powers. *IEEE Trans. Inform. Theory*, 11(2):267–271, April 1965.

[8] S.G. Bobkov. Upper bounds for Fisher information. *Electron. J. Probab.*, 27:1–44, 2022.

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

[10] E.A. Carlen and A. Soffer. Entropy production by block variable summation and central limit theorems. *Comm. Math. Phys.*, 140(2):339–371, September 1991.

[11] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In R.C. Gunning, editor, *Problems in Analysis*, pages 195–200. Princeton University Press, Princeton, NJ, 1971.

[12] T.A. Courtade. Monotonicity of entropy and Fisher information: A quick proof via maximal correlation. *Commun. Inf. Syst.*, 16(2):111–115, Januray 2016.

[13] T.A. Courtade, M. Fathi, and A. Pananjady. Quantitative stability of the entropy power inequality. *IEEE Trans. Inform. Theory*, 64(8):5691–5703, August 2018.

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

[15] R.M. Dudley. *Real analysis and probability*. Cambridge University Press, Cambridge, U.K., second edition, 2002.

[16] R. Eldan and D. Mikulincer. Stability of the Shannon-Stam inequality via the Föllmer process. *Probab. Theory Related Fields*, 177(3-4):891–922, 2020.

[17] M. Fradelizi, L. Gavalakis, and M. Rapaport. On the monotonicity of discrete entropy for log-concave random vectors on $\mathbb{Z}^d$. *arXiv e-prints*, `2401.15462 [math.PR]`, January 2024.

[18] L. Gavalakis. Approximate discrete entropy monotonicity for log-concave sums. *Comb., Probab. Comput.*, 33(2):196–209, March 2024.

[19] L. Gavalakis and Kontoyiannis. Conditions for equality and stability in Shannon's and Tao's entropy power inequalities. *arXiv e-prints*, `2509.14021 [math.PR]`, September 2025.

[20] L. Gavalakis and I. Kontoyiannis. Entropy and the discrete central limit theorem. *Stoch. Proc. Appl.*, 170:104294, June 2024.

[21] L. Gavalakis, I. Kontoyiannis, and M. Madiman. The entropic doubling constant and robustness of Gaussian codebooks for additive-noise channels. *IEEE Trans. Inform. Theory*, 70(12):8467–8477, December 2024.

[22] W.T. Gowers, B. Green, F. Manners, and T. Tao. On a conjecture of Marton. *Ann. of Math.*, 201(2):515–549, March 2025.

[23] B. Green, F. Manners, and T. Tao. Sumsets and entropy revisited. *Random Struct. Algorithms*, 66(1):e21252, 2025.

[24] Q. Guan. A note on Bourgain's slicing problem. *arXiv e-prints*, `2412.09075 [math.MG]`, December 2024.

[25] O. Johnson and A.R. Barron. Fisher information inequalities and the central limit theorem. *Probab. Theory Related Fields*, 129(3):391–409, July 2004.

[26] R. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.*, 13:541–559, 1995.

[27] C.A.J. Klaassen. On an inequality of Chernoff. *Ann. Probab.*, 13(3):966–974, August 1985.

[28] B. Klartag. Logarithmic bounds for isoperimetry and slices of convex sets. *Ars Inven. Anal.*, pages Paper No. 4, 17, 2023.

[29] B. Klartag and J. Lehec. Affirmative resolution of Bourgain's slicing problem using Guan's bound. *arXiv e-prints*, `2412.15044 [math.MG]`, December 2024.

[30] B. Klartag and O. Ordentlich. The strong data processing inequality under the heat flow. *IEEE Trans. Inform. Theory*, 71(5):3317–3333, May 2025.

[31] I. Kontoyiannis and M. Madiman. Sumset and inverse sumset inequalities for differential entropy and mutual information. *IEEE Trans. Inform. Theory*, 60(8):4503–4514, August 2014.

[32] J. Lehec. Representation formula for the entropy and functional inequalities. *Ann. Inst. Henri Poincaré*, 49(3):885–899, August 2013.

[33] R. Li, L. Gavalakis, and Kontoyiannis. Entropic additive energy and entropy inequalities for sums and products. *arXiv e-prints*, `2506.20813 [cs.IT]`, June 2025.

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

[35] M. Madiman and A.R Barron. Generalized entropy power inequalities and monotonicity properties of information. *IEEE Trans. Inform. Theory*, 53(7):2317–2329, July 2007.

[36] M. Madiman and I. Kontoyiannis. Entropy bounds on abelian groups and the Ruzsa divergence. *IEEE Trans. Inform. Theory*, 64(1):77–92, January 2018.

[37] R. Raghavan. Improved Bounds for the Freiman-Ruzsa Theorem. *arXiv preprint arXiv:2512.11217,* 2025.

[38] O. Rioul. A simple proof of the entropy-power inequality via properties of mutual information. In *2007 IEEE International Symposium on Information Theory (ISIT)*, pages 46–50, Nice, France, June 2007.

[39] O. Rioul. Information theoretic proofs of entropy power inequalities. *IEEE Trans. Inform. Theory*, 57(1):33–55, January 2011.

[40] O. Rioul. Optimal transportation to the entropy-power inequality. In *Workshop on Information Theory and Applications (ITA)*, UCSD, San Diego, CA, February 2017.

[41] O. Rioul. Yet another proof of the entropy power inequality. *IEEE Trans. Inform. Theory*, 63(6):3595–3599, June 2017.

[42] C.E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27(3):379–423, 623–656, 1948.

[43] A.J. Stam. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Contr.*, 2(2):101–112, 1959.

[44] S.J. Szarek and D. Voiculescu. Shannon's entropy power inequality via restricted Minkowski sums. In V.D. Milman and G. Schechtman, editors, *Geometric Aspects of Functional Analysis: Israel Seminar 1996-2000*, pages 257–262. Springer, Berlin, Heidelberg, 2000.

[45] T. Tao. Sumset and inverse sumset theory for Shannon entropy. *Comb., Probab. Comput.*, 19(4):603–639, July 2010.

[46] S. Verdú and D. Guo. A simple proof of the entropy-power inequality. *IEEE Trans. Inform. Theory*, 52(5):2165–2166, May 2006.