A dimensional improvement of the Entropy Power Inequality under marginal assumptions

Matthieu Fradelizi*, Lampros Gavalakis†,1, Martin Rapaport‡

* ‡ Univ Gustave Eiffel, Univ Paris Est Creteil,
CNRS, LAMA UMR8050 F-77447
Marne-la-Vallée, France
† University of Cambridge, DPMMS, UK
‡ Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, 15213, PA, United States

*matthieu.fradelizi@univ-eiffel.fr, †lg560@cam.ac.uk, ‡mrapapor@andrew.cmu.edu

***Abstract*—We show that the Entropy Power Inequality can be improved in dimension $n > 1$ under the assumption that the two random variables have the same entropy along some marginal. This is an analogous improvement to Bergström's inequality for determinants and Bonnesen's inequality for volumes. In fact we establish a more general inequality, whose proof generalizes an information-theoretic proof of Bergström's inequality due to Dembo, Cover and Thomas (1991). Moreover, we characterize the equality case in our main inequality. Finally, we provide a similar inequality for the Fisher information.**

A full version of this paper is available as: arXiv:2501.10309.

I. Introduction

The differential entropy of a Gaussian can be expressed in terms of the logarithm of the determinant of the covariance matrix. Thus, several properties of the entropy transfer to logarithms of determinants of symmetric, positive semi-definite matrices and a number of determinant inequalities may be proved using entropy [1], [2].

Furthermore, the exponential of the entropy can be thought of as the volume of the typical set. Therefore, entropy behaves in many ways similar to volume. This analogy between entropy and volume has been observed at least as early as in the work of Costa and Cover [3], where the connection between the Entropy Power Inequality (EPI) and the Brunn-Minkowski inequality is examined (see [1] for an extensive review and a common proof of these inequalities).

Motivated by these connections, we will prove analogues of the Bergström and Bonnesen inequalities for entropy and Fisher information, which can be seen as refinements of certain inequalities for determinants and volumes respectively. The classical matrix form of the Bergström inequality is the following:

Theorem 1: [4] Let A and B be two $n \times n$ positive definite real symmetric matrices, and denote by A_i and B_i the two $(n-1)\times(n-1)$ matrices resulting from A and B by deleting the i-th row and the i-th column. Then we have

$$\frac{\det(A+B)}{\det(A_i+B_i)} \geq \frac{\det(A)}{\det(A_i)} + \frac{\det(B)}{\det(B_i)},$$

for every $i \in \{1,\dots,n\}$.

The above theorem is, by setting $A = \lambda S$ and $B = (1-\lambda)T$, equivalent to the statement that $\frac{\det(A)}{\det(A_i)}$ is concave in A. A proof of this statement using entropy was given in [1], by considering the entropy of Gaussian random vectors with covariances A and B and observing that conditioning reduces entropy.

The first question that motivated the current work is whether an entropic analogue of Theorem 1 can be obtained. That is, can we derive an inequality is true for any random variables and generalizes the inequality obtained in the proof of [1, Theorem 30] for Gaussians? Our first main result, Theorem 4 provides such an inequality (inequality (7)).

The concavity of $\frac{\det(A)}{\det(A_i)}$ directly implies that if $\det(A_i) = \det(B_i)$ for some i, then

$$\det(\lambda A + (1-\lambda)B)) \geq \lambda \det(A) + (1-\lambda)\det(B). \quad (1)$$

It is well known that $A \mapsto \det(A)^{\frac{1}{n}}$ is a concave functional of the matrix A, i.e. $\det(\lambda A + (1-\lambda)B)^{\frac{1}{n}} \geq \lambda \det(A)^{\frac{1}{n}} + (1-\lambda)\det(B)^{\frac{1}{n}}$. Inequality (1) is an improvement of the latter (by concavity of $x \to x^{\frac{1}{n}}$) and can fail in general without the assumption that the matrices obtained by removing some column and the corresponding row have equal determinants.

Another theorem due to Bonnesen implies linear refinements of the Brunn-Minkowski inequality and can be seen as a weaker volume-analogue of (1). Here and in what follows, we denote with $|K|_n$ the volume of a set K in $\mathbb{R}^n$.

Theorem 2: [5] Let A and B be convex bodies in $\mathbb{R}^n$ and let $\theta \in S^{n-1}$ then

$$\frac{|A+B|_n}{\left(|P_{\theta^\perp}A|_{n-1}^{\frac{1}{n-1}} + |P_{\theta^\perp}B|_{n-1}^{\frac{1}{n-1}}\right)^{n-1}} \geq \frac{|A|_n}{|P_{\theta^\perp}A|_{n-1}} + \frac{|B|_n}{|P_{\theta^\perp}B|_{n-1}},$$

[1]L.G. was supported by the Economic and Physical Sciences Research Council [Grant Ref:EP/Y028732/1].

where $P_\theta^\perp A$ denotes the orthogonal projection of A onto the hyperplane with normal vector θ.

Let K and L be two convex bodies in $\mathbb{R}^n$ such that $|P_{\theta^\perp} K|_{n-1} = |P_{\theta^\perp} L|_{n-1}$ for some $\theta \in S^{n-1}$. Then, applying Bonnesen's inequality above to $A = (1-\lambda)K$ and $B = \lambda L$ we deduce that

$$|(1-\lambda)K + \lambda L|_n \geq (1-\lambda)|K|_n + \lambda |L|_n, \qquad (2)$$

which is the linear improvement of Brunn-Minkowski's inequality. It would be natural to conjecture that a stronger Bonnesen's inequality in the form

$$\frac{|A+B|_n}{|P_{\theta^\perp}(A+B)|_{n-1}} \geq \frac{|A|_n}{|P_{\theta^\perp} A|_{n-1}} + \frac{|B|_n}{|P_{\theta^\perp} B|_{n-1}} \qquad (3)$$

holds for any convex bodies A and B in $\mathbb{R}^n$ but, in [6], it was disproved in general for $n \geq 3$ and proved for A and B being zonoids in $\mathbb{R}^3$.

In view of the improvements (1) and (2) and the close connection between the Brunn-Minkowski and EPI, it is natural to ask under which assumptions one can improve the EPI in an analogous manner.

Recall that the differential entropy $h(X)$ and the entropy power $N(X)$ of a random vector X in $\mathbb{R}^n$ with density f are defined as

$$h(X) = -\int_{\mathbb{R}^n} f(x)\log f(x)dx \quad \text{and} \quad N(X) = e^{\frac{2}{n}h(X)},$$

whenever the integral makes sense and we put conventionally $h(X) = -\infty$ otherwise.

Let X and Y be two independent random variables in $\mathbb{R}^d$ such that the entropies $h(X), h(Y)$ and $h(X+Y)$ exist. One of the equivalent formulations of the EPI states that

$$N(X+Y) \geq N(X) + N(Y),$$

i.e.

$$e^{\frac{2}{n}h(X+Y)} \geq e^{\frac{2}{n}h(X)} + e^{\frac{2}{n}h(Y)}.$$

In Corollary 6 we show that if two n-dimensional random vectors have some $n-1$-dimensional marginal with the same entropy, then the EPI can be improved by removing the $\frac{1}{n}$ factor from the exponent.

Moreover, we characterize the equality case, which turns out to be if and only if X, Y are Gaussians with the same covariance matrix up to the last element of the diagonal. As expected, the latter condition on the covariance matrices is the same as in the equality case in (1) (see [7]).

The Fisher information of X is defined as

$$I(X) = \int_{\mathbb{R}^n} \frac{\|\nabla f(x)\|^2}{f(x)}\,dx,$$

when the integral is well-defined and $I(X) = \infty$ otherwise. As an analogue to the Fisher information inequality

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

which holds for any independent X, Y, Dembo, Cover, Thomas [1] asked whether the inequality

$$\frac{|K+L|}{|\partial(K+L)|} \geq \frac{|K|}{|\partial K|} + \frac{|L|}{|\partial L|} \qquad (4)$$

holds true for convex bodies. More generally, it was asked by Millman, for which values of k is

$$\frac{V_k(K+L)}{V_{k-1}(K+L)} \geq \frac{V_k(K)}{V_{k-1}(K)} + \frac{V_k(L)}{V_{k-1}(L)}, \qquad (5)$$

where $V_k(K)$ denote the k-th mixed volumes, i.e. the coefficient of t^{n-k} in the polynomial expansion of $|K + tB_2^n|$ in $t > 0$ where B_2^n is the Euclidean ball in $\mathbb{R}^n$. Inequality (5) was shown to hold true [8] if and only if $k = 1, 2$, implying that (4) is not true in general.

In view of the analogy between entropy and volume, our entropic Bergström inequality (7) also serves as an entropic analogue of (4). Moreover, it is in the analogous form of the volume inequality (3), which is not true in general, i.e. the entropic analogue is always true even though the volume version fails in general. This should not be surprising, since entropy usually serves as a more flexible analogue of volume. This is often the case in discrete settings as well; it was recently highlighted by the breakthrough proof of Marton's conjecture by Gowers, Green, Manners and Tao [9]–[11], where the observation that entropy behaves well under group homomorphisms, in contrast to cardinality, turned out to be crucial.

Outline. In Section II we prove and discuss our main inequalities for entropy powers.

As a corollary of our main result we also obtain an analogous improvement of the isoperimetric inequality for entropies in Section III.

Finally, in view of the different forms of the Bergström and Bonnesen inequalities, we establish a different analogue for the Fisher information. To that end, in Section IV we define a conditional version of the Fisher information and prove an inequality, which resembles Bergström's inequality. This turns out to be stronger than (in the sense that it implies) the convolution inequality for Fisher information, sometimes referred to as Blachman-Stam inequality. We do not know whether our Fisher information inequality implies our entropic Bergström inequality.

Notation. We write capital letters X, Y for random variables (resp. vectors) and small letters x, y for specific realizations of these. If $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ we write $X^{n-1} := (X_1, \ldots, X_{n-1})$ to denote the first $n-1$ coordinates. When it is not clear from the context, we will write $h_n(X)$ and $N_n(X)$ for the entropy and entropy power respectively, to emphasize that the integral in the definition of entropy (see Section II) is with respect to the Lebesgue measure in $\mathbb{R}^n$.

II. Entropy inequalities

In the proof of our main result we need a conditional form of the EPI, the proof of which is essentially due to Stam [12], although we express it in a slightly more general setting (see also Bergmans [13, Lemma II] and [14] where quantum versions are of interest).

Let X, Z, Y be three random vectors with densities. We say that $X \to Z \to Y$ form a Markov chain if, given on

Z, X and Y are conditionally independent, i.e. for a.e. z, $f(x,y|z) = f(x|z)f(y|z)$, where $f(x,y,z)$ is the joint density of X, Y, Z and for any z, $f(x,y|z) = \frac{f(x,y,z)}{f(z)} = \frac{f(x,y,z)}{\int f(u,v,z)dudv}$ is the conditional density of (X,Y) given $Z = z$, defined for a.e. z, and analogously $f(x|z) = \frac{f(x,z)}{f(z)} = \frac{\int f(x,v,z)dv}{\int f(u,v,z)dudv}$.

Lemma 3 ([12], [13]): Suppose $X, Y \in \mathbb{R}^n$ and Z taking with values in some space Ω are such that $X \to Z \to Y$ form a Markov chain and, given Z, X and Y have conditional densities on $\mathbb{R}^n$. Then

$$N(X+Y|Z) \geq N(X|Z) + N(Y|Z),$$

where for any random vectors U, V such that U has a conditional density given V in $\mathbb{R}^n$, $N(U|V) := e^{\frac{2}{n}h(U|V)}$. Moreover, there is equality, if and only if for almost every z, the conditional densities given $Z = z$ of X and Y are Gaussian with proportional covariances.

Proof: See Stam [12]. The equality case follows from the equality case in this form of the EPI [1], $\tilde{X}|_{Z=z}$ and $\tilde{Y}|_{Z=z}$ are Gaussian with the same covariance matrix. By the definitions of $\tilde{X}$ and $\tilde{Y}$, X, Y are also Gaussian having as covariance matrix each a different multiple of the common covariance of $\tilde{X}$ and $\tilde{Y}$. □

Theorem 4 (Entropy analogue of Bergström's inequality): Let $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ be two independent random vectors in $\mathbb{R}^n$. Let $X^{n-1} = (X_1, \ldots, X_{n-1})$ and $Y^{n-1} = (Y_1, \ldots, Y_{n-1})$. Then, the following inequalities hold true and are equivalent:

$$\begin{aligned} &e^{2h(\sqrt{1-\lambda}X_n + \sqrt{\lambda}Y_n | \sqrt{1-\lambda}X^{n-1} + \sqrt{\lambda}Y^{n-1})} \\ &\geq (1-\lambda)e^{2h(X_n|X^{n-1})} + \lambda e^{2h(Y_n|Y^{n-1})} \end{aligned} \tag{6}$$

and

$$\begin{aligned} &\frac{N(X+Y)^n}{N_{n-1}(X^{n-1}+Y^{n-1})^{n-1}} \\ &\geq \frac{N(X)^n}{N_{n-1}(X^{n-1})^{n-1}} + \frac{N(Y)^n}{N_{n-1}(Y^{n-1})^{n-1}}. \end{aligned} \tag{7}$$

Proof: Inequality (6) can be seen to be equivalent to

$$e^{2h(X_n+Y_n|X^{n-1}+Y^{n-1})} \geq e^{2h(X_n|X^{n-1})} + e^{2h(Y_n|Y^{n-1})}$$

by setting $\tilde{X} = \sqrt{1-\lambda}X$ (and analogously for Y) and scaling.

The latter is equivalent to (7) by the definition of entropy power and the chain rule for differential entropy.

We prove (6). Since conditioning reduces entropy,

$$\begin{aligned} &h\left(\sqrt{1-\lambda}X_n + \sqrt{\lambda}Y_n \mid \sqrt{1-\lambda}X^{n-1} + \sqrt{\lambda}Y^{n-1}\right) \\ &\geq h\left(\sqrt{1-\lambda}X_n + \sqrt{\lambda}Y_n \mid X^{n-1}, Y^{n-1}\right). \end{aligned}$$

Thus,

$$\begin{aligned} &e^{2h(\sqrt{1-\lambda}X_n + \sqrt{\lambda}Y_n | \sqrt{1-\lambda}X^{n-1} + \sqrt{\lambda}Y^{n-1})} \\ &\geq e^{2h\left(\sqrt{1-\lambda}X_n + \sqrt{\lambda}Y_n | X^{n-1}, Y^{n-1}\right)} \\ &\geq (1-\lambda)e^{2h(X_n|X^{n-1},Y^{n-1})} + \lambda e^{2h(Y_n|X^{n-1},Y^{n-1})} \qquad (8) \\ &= (1-\lambda)e^{2h(X_n|X^{n-1})} + \lambda e^{2h(Y_n|Y^{n-1})}, \end{aligned}$$

where in (8) we have used the conditional EPI, Lemma 3, with $Z = (X^{n-1}, Y^{n-1})$, after noting that by the independence of X and Y, X_n and Y_n are conditionally independent given (X^{n-1}, Y^{n-1}).

□

Remark 5:

1) Defining the map

$$f : \lambda \mapsto \frac{N\left(\sqrt{\lambda}X + \sqrt{1-\lambda}Y\right)^n}{N_{n-1}\left(\sqrt{\lambda}X^{n-1} + \sqrt{1-\lambda}Y^{n-1}\right)^{n-1}}, \tag{9}$$

it can be seen from the proof of Theorem 4 that the following form is also equivalent to (7):

$$\begin{aligned} f(\lambda) &\geq \frac{\lambda N(X)^n}{N_{n-1}(X^{n-1})^{n-1}} + \frac{(1-\lambda)N(Y)^n}{N_{n-1}(Y^{n-1})^{n-1}} \qquad (10) \\ &= \lambda f(1) + (1-\lambda)f(0). \qquad (11) \end{aligned}$$

Ball, Nayar and Tkocz [15] conjectured that the map $\lambda \to h(\sqrt{1-\lambda}X + \sqrt{\lambda}Y)$ is concave for i.i.d. log-concave X, Y in dimension 1. A partial result in a different direction was given in [16], [17], where it was shown that this entropy-map is concave for Gaussian (scale) mixtures.

In view of the above conjecture and the partial answer for Gaussian mixtures it is natural to ask for which class of random variables (if any) is the map defined in (9) concave.

Note that, in contrast to these works, when restricting to $n = 1$, we are asking for concavity of the entropy power, rather than the entropy itself, which is stronger, since concavity implies log-concavity.

When one of the two random vectors is Gaussian, by homogeneity, concavity of the map defined above is equivalent to concavity of

$$f : \lambda \mapsto \frac{N\left(X + \sqrt{\lambda}Z\right)^n}{\left(N_{n-1}(X^{n-1} + \sqrt{\lambda}Z^{n-1})\right)^{n-1}},$$

analogously to [18, Proposition 2.2]. The latter is known to hold true in dimension 1 [19].

2) Inequality (7) implies the EPI, $N(X+Y) \geq N(X) + N(Y)$, for $n > 1$, in the sense that the latter can be deduced from the former via the following inductive argument:

Assume (7) and the EPI for $n = 1$. By the EPI for $n-1$, $N_{n-1}(X^{n-1}+Y^{n-1}) \geq N_{n-1}(X^{n-1}) + N_{n-1}(Y^{n-1})$, and therefore (7) implies

$$\begin{aligned} N(X+Y)^n &\geq \theta\frac{N(X)^n}{\theta^n} + (1-\theta)\frac{N(Y)^n}{(1-\theta)^n} \\ &\geq (N(X) + N(Y))^n, \end{aligned}$$

where $\theta = \frac{N_{n-1}(X^{n-1})}{N_{n-1}(X^{n-1}) + N_{n-1}(Y^{n-1})}$, by convexity of $x \to x^n, x \geq 0$, and the EPI follows. Note that in the proof we only used the EPI in $n = 1$.

3) The steps in the proof of Theorem 4 are generalizations of the proof of [1], which restricted to Gaussian random variables. Taking X, Y Gaussian in our result, we recover Bergström's inequality, Theorem 1, for determinants.
4) Although we state Theorem 4 for simplicity using $N_{n-1}(X^{n-1})$ and $N_{n-1}(Y^{n-1})$, it can be generalized, via a change of axes, in the sense that the same conclusion holds with $N_{n-1}(AX)$ and $N_{n-1}(AY)$ instead, for any projection $A : \mathbb{R}^n \to \mathbb{R}^{n-1}$.
5) The same proof also works to obtain the following entropic analogue of Ky-Fan's Theorem [20], which generalizes (6): for $X = (X_1, \dots, X_n)$ and $I \subset \{1, \dots, n\}$, write $X_I = \{X_i\}_{i \in I}$. For I with $|I| = k$, we have
$$e^{\frac{2}{k} h(\sqrt{1-\lambda} X_I + \sqrt{\lambda} Y_I | \sqrt{1-\lambda} X_{I^c} + \sqrt{\lambda} Y_{I^c})} \geq (1-\lambda) e^{\frac{2}{k} h(X_I | X_{I^c})} + \lambda e^{\frac{2}{k} h(Y_I | Y_{I^c})}.$$

Theorem 4 implies an entropic version of Bonnesen's inequality (2), Corollary 6 below. In view of the concavity of $x \to x^{\frac{1}{n}}, x \geq 0$, it shows that, under the assumption that some $n-1$–dimensional subvectors of X and Y have the same entropy, the EPI can be improved. Again, although we state it for simplicity under the assumption $h(X^{n-1}) = h(Y^{n-1})$, the latter can be relaxed via a change of axes to $h(AX) = h(AY)$ for some projection $A : \mathbb{R}^n \to \mathbb{R}^{n-1}$.

Corollary 6 (Improved EPI under marginal assumptions): Let $X = (X_1, \dots, X_n)$ and $Y = (Y_1, \dots, Y_n)$ be two independent random vectors in $\mathbb{R}^n$. Let $X^{n-1} = (X_1, \dots, X_{n-1})$ and $Y^{n-1} = (Y_1, \dots, Y_{n-1})$ and suppose that $h(X^{n-1}) = h(Y^{n-1})$. Then, for all $\lambda \in [0,1]$,
$$e^{2h(\sqrt{1-\lambda} X + \sqrt{\lambda} Y)} \geq (1-\lambda) e^{2h(X)} + \lambda e^{2h(Y)}. \tag{12}$$
Moreover, there is equality if and only if X and Y are Gaussians having the same covariance matrix except the last element of the diagonal.

Proof: This follows by applying the EPI on the denominator of the left-hand side in (10) and simplifying, since $N_{n-1}(X^{n-1}) = N_{n-1}(Y^{n-1})$.

For the equality case, note that if there is equality in (12), all inequalities used in its proof have to be equalities. Thus, the use of the EPI suggests that X^{n-1} and Y^{n-1} are Gaussians with proportional covariances. Since, by assumption, $h(X^{n-1}) = h(Y^{n-1})$, the covariances of X^{n-1} and Y^{n-1} have in addition the same determinant. Therefore, they have to be equal to each other.

Moreover, we must have equality in (8). By the equality case in Lemma 3 the conditional distributions X_n and Y_n given X^{n-1} and Y^{n-1} respectively, are Gaussian. Therefore, X and Y are both Gaussian.

The fact that the elements of the last rows (respectively columns) of the covariances are the same follows from linear algebraic considerations. We skip the details and refer to the full version of the paper.

□

III. Isoperimetric Inequalities

The isoperimetric inequality for entropies asserts that
$$I(X)N(X) \geq 2\pi e n, \tag{13}$$
is due to Stam [12] (see also [1, Theorem 16]) and is a consequence of the EPI. By our strengthened EPI, we obtain a sharpening of the isoperimetric inequality, Corollary 7 below.

Corollary 7: Let X be a random vector in $\mathbb{R}^n$. Then, the following inequality is satisfied:
$$\frac{I(X)N(X)}{2\pi e} \geq \left(\frac{N(X^{n-1})}{N(X)} \right)^{n-1} + \frac{(n-1)N(X)}{N_{n-1}(X^{n-1})}. \tag{14}$$

Proof: Let X be a random vector in $\mathbb{R}^n$, Z be a standard Gaussian vector independent of X and note that $N(Z) = 2\pi e$. By Theorem 4 and a first-order Taylor expansion of the convex function $t \mapsto (N_{n-1}(X^{n-1}) + tN_{n-1}(Z^{n-1}))^{n-1}, t \geq 0$, we obtain
$$\begin{aligned} &N(X + \sqrt{t}Z)^n \\ &\geq \left(N_{n-1}(X^{n-1})^{n-1} + 2\pi e t(n-1) N_{n-1}(X^{n-1})^{n-2} + o(t)\right) \\ &\times \left(\frac{N(X)^n}{N_{n-1}(X^{n-1})^{n-1}} + 2\pi e t \right). \end{aligned}$$
Therefore,
$$\begin{aligned} &N(X + \sqrt{t}Z)^n \geq N(X)^n \\ &+ 2\pi e t \left(N_{n-1}(X^{n-1})^{n-1} + (n-1) \frac{N(X)^n}{N_{n-1}(X^{n-1})} \right) + o(t). \end{aligned} \tag{15}$$
This implies, as $t \to 0$,
$$\begin{aligned} &\lim_{t \to 0} \frac{N(X + \sqrt{t}Z)^n - N(X)^n}{t} \\ &\geq 2\pi e \left(N_{n-1}(X^{n-1})^{n-1} + (n-1) \frac{N(X)^n}{N_{n-1}(X^{n-1})} \right). \end{aligned}$$
We recall de Bruijn's identity [1, Theorem 14]
$$\frac{d}{dt} h(X + \sqrt{t}Z) = \frac{1}{2} I(X + \sqrt{t}Z). \tag{16}$$
Therefore the term on the left-hand side of (15) corresponds to $\frac{d}{dt} N(X + \sqrt{t}Z)^n \mid_{t=0}$, which by (16) equals $I(X)N(X)^n$ and the claimed inequality follows.

□

To see that (14) is indeed stronger than the usual isoperimetric inequality for entropies, note that by letting $a = \frac{N(X^{n-1})}{N(X)}$ in (14) and noting that the arithmetic-geometric mean inequality implies $\frac{1}{n} a^{n-1} + \frac{n-1}{n} \frac{1}{a} \geq 1$, the isoperimetric inequality (13) follows.

A result of similar spirit was proved by Courtade [21, Theorem 4], which can be seen as an improvement of the isoperimetric inequality of different kind, namely under an entropy jump assumption on self-convolution.

IV. Fisher information inequalities

In what follows we write $I(f) = I(X)$ for the Fisher information of X when the latter has density f. We recall that a sufficient condition for the Fisher information, defined in the Introduction, to be well-defined and finite is that $\sqrt{f}$ belongs to the Sobolev space $W_1^2(\mathbb{R}^n)$ [22].

The Blachman-Stam or Fisher information inequality, asserts that for any pair of independent random vectors X and Y in $\mathbb{R}^n$,

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

We define

$$I_{P_n}(X) := \int_{\mathbb{R}^n} \frac{\langle \nabla f, e_n \rangle^2}{f(x)}\, dx, \tag{17}$$

whenever the integral is well-defined and we call I_{P_n} the *projective Fisher information* of X.

*Definition 8 (*Conditional Fisher Information): For two random vectors $X \in \mathbb{R}^n, Y \in \mathbb{R}^m$, such that the conditional density of X given Y, say $f_{X|Y}(\cdot|\cdot)$, exists a.s. with respect to the Lebesgue measure on $\mathbb{R}^k$ for some $1 \leq k \leq n$, we define the conditional Fisher Information of X given Y as the expected Fisher Information of the conditional density, that is

$$I(X|Y) := \int_{\mathbb{R}^m} f_Y(y) I\big(f_{X|Y}(\cdot|y)\big) dy,$$

whenever the integral is well-defined, and the Fisher Information of the conditional distribution is to be understood as an integral in $\mathbb{R}^k$.

Two properties of the projective and conditional Fisher information of a random vector $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ are straightforward from the definitions:

1)
$$I_{P_n}(X) = I(X_n|X^{n-1}). \tag{18}$$
This is just by expanding the joint density as a product of the marginal and the conditional density and using Fubini's theorem.
2) By (18) and since for any vector u, $\langle u, e_n\rangle^2 \leq \|u\|^2$, we have,
$$I(X_n|X^{n-1}) = I_{P_n}(X) \leq I(X). \tag{19}$$

Applying the Fisher information inequality to a particular operator, we obtain the following conditional Fisher information inequality:

Theorem 9: Let X, Y be two random vectors in $\mathbb{R}^n$ with finite Fisher informations. Then

$$I_{P_n}(X+Y)^{-1} \geq I_{P_n}(X)^{-1} + I_{P_n}(Y)^{-1}, \tag{20}$$

or, equivalently,

$$\begin{aligned} &I(X_n + Y_n|X^{n-1} + Y^{n-1})^{-1} \\ &\geq I(X_n|X^{n-1})^{-1} + I(Y_n|Y^{n-1})^{-1}, \end{aligned}$$

provided that the conditional Fisher informations exist.

Proof: Let $T_m : \mathbb{R}^n \to \mathbb{R}^n$ be the linear operator

$$T_m = \mathrm{I}_n + \left(\frac{1}{m} - 1\right) e_n e_n^T,$$

where I_n denotes the identity matrix. Then, writing f for the density of X, $T_m(X)$ has density

$$f_{T_m}(x) := \frac{1}{|\det T_m|} f(T_m^{-1}x), x \in \mathbb{R}^n.$$

Thus, the Fisher information $I(T_m(X))$ is given by

$$\begin{aligned} I(T_m(X)) &= \int_{\mathbb{R}^n} \frac{\|\nabla f_{T_m}(x)\|^2}{f_{T_m(x)}}\, dx = \int_{\mathbb{R}^n} \frac{\|(T_m^{-1})^T \nabla f(y)\|^2}{f(y)}\, dy \\ &= \int_{\mathbb{R}^n} \frac{\sum_{i=1}^{n-1} (\frac{\partial}{\partial y_i} f)^2 + m^2 (\frac{\partial}{\partial y_n} f)^2}{f(y)}\, dy. \end{aligned}$$

Therefore, using the finiteness of the Fisher information of X,

$$\lim_{m\to\infty} \frac{I(T_m(X))}{m^2} = \int_{\mathbb{R}^n} \frac{(\frac{\partial}{\partial y_n} f)^2}{f(y)}\, dy = I_{P_n}(X).$$

Analogously, $\lim_{m\to\infty} \frac{I(T_m(Y))}{m^2} = I_{P_n}(Y)$ and $\lim_{m\to\infty} \frac{I(T_m(X+Y))}{m^2} = I_{P_n}(X+Y)$.

Now the result follows by applying the Blachman-Stam inequality to the independent random variables $T_m(X)$ and $T_m(Y)$, after dividing by m^2 and letting m tend to infinity. □

Remark 10: An examination of the proof shows that the same inequality can be obtained for the more general functional $I_{P^u} := \int_{\mathbb{R}^n} \frac{\langle \nabla f, u\rangle^2}{f(x)} dx, u \in \mathbb{S}^{n-1}$. Using the identity

$$\int_{\mathbb{S}^{n-1}} \langle u, v\rangle^2 d\sigma(u) = \frac{\|v\|^2}{n}, \tag{21}$$

where $d\sigma$ denotes the uniform measure on the sphere and which holds for any vector v, we recover the Blachman-Stam inequality. Indeed, by Minkowski's inequality for $p = -1$,

$$\begin{aligned} I(X+Y) &= n \int_{\mathbb{S}^{n-1}} I_{P^u}(X+Y))\, d\sigma(u) \\ &\leq n \int_{\mathbb{S}^{n-1}} \big(I_{P^u}(X))^{-1} + I_{P^u}(Y))^{-1}\big)^{-1}\, d\sigma(u) \\ &\leq \Big(I(X)^{-1} + I(Y^{-1})\Big)^{-1}. \end{aligned}$$

V. Conclusion

Inspired by classical inequalities for determinants and volumes, we established an entropic inequality that strengthens the EPI when two independent random vectors share a common marginal entropy. As a corollary, we derived an improved isoperimetric inequality for entropy. In addition, motivated by the projective nature of our entropic inequality and the quest to fully understand the role of Fisher information in this analogy, we proposed a conditional version of the Fisher information and established a convolution inequality.

The connection between entropy and volume, although very fruitful, does not yield a one-to-one correspondence. One task that remains elusive is to identify an appropriate analogue of mixed volumes. We believe that the results of the present work and especially (7) could hind towards possible entropic counterparts. This will be the subject of future work.

References

[1] A. Dembo, T. M. Cover, and J. A. Thomas, "Information theoretic inequalities," *IEEE Transactions on Information theory*, vol. 37, no. 6, pp. 1501–1518, 1991.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. New York, NY, USA: Wiley-Interscience, 2006.

[3] M. Costa and T. Cover, "On the similarity of the entropy power inequality and the Brunn-Minkowski inequality (," *IEEE Transactions on Information Theory*, vol. 30, no. 6, pp. 837–839, 1984.

[4] H. Bergström, "A triangle inequality for matrices," *Den Elfte Skandinaviske Matematikerkongress*, pp. 264–267, 1949.

[5] T. Bonnesen and W. Fenchel, "Konvexe körper," *New York*, 1948.

[6] M. Fradelizi, M. Madiman, M. Meyer, and A. Zvavitch, "On the volume of the Minkowski sum of zonoids," *J. Funct. Anal.*, vol. 286, no. 3, pp. Paper No. 110 247, 41, 2024. [Online]. Available: https://doi.org/10.1016/j.jfa.2023.110247

[7] N. Lombardi and E. Saorín Gómez, "Short note on some geometric inequalities derived from matrix inequalities," *Positivity*, vol. 28, no. 2, p. 26, 2024.

[8] M. Fradelizi, M. Meyer, and A. Giannopoulos, "Some inequalities about mixed volumes," *Israel Journal of Mathematics*, vol. 135, pp. 157–179, 2003.

[9] W. Gowers, B. Green, F. Manners, and T. Tao, "On a conjecture of Marton," *arXiv preprint arXiv:2311.05762*, 2023.

[10] ——, "Marton's conjecture in abelian groups with bounded torsion," *arXiv preprint arXiv:2404.02244*, 2024.

[11] B. Green, F. Manners, and T. Tao, "Sumsets and entropy revisited," *Random Structures & Algorithms*, vol. 66, no. 1, p. e21252, 2025.

[12] A. J. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Information and Control*, vol. 2, no. 2, pp. 101–112, 1959.

[13] P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 279–280, 1974.

[14] G. De Palma and S. Huber, "The conditional entropy power inequality for quantum additive noise channels," *Journal of Mathematical Physics*, vol. 59, no. 12, 2018.

[15] K. Ball, P. Nayar, and T. Tkocz, "A reverse entropy power inequality for log-concave random vectors," *Studia Mathematica*, vol. 235, pp. 17–30, 2016.

[16] A. Eskenazis and L. Gavalakis, "On the entropy and information of gaussian mixtures," *Mathematika*, vol. 70, no. 2, p. e12246, 2024.

[17] ——, "Gaussian mixtures: Convexity properties and CLT rates for the entropy and fisher information," in *2024 IEEE International Symposium on Information Theory (ISIT)*, 2024, pp. 3594–3599.

[18] M. Fradelizi and A. Marsiglietti, "On the analogue of the concavity of entropy power in the Brunn–Minkowski theory," *Advances in Applied Mathematics*, vol. 57, pp. 1–20, 2014.

[19] M. Costa, "A new entropy power inequality," *IEEE Transactions on Information Theory*, vol. 31, no. 6, pp. 751–760, 1985.

[20] K. Fan, "Some inequalities concerning positive-definite hermitian matrices," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 51, no. 3. Cambridge University Press, 1955, pp. 414–421.

[21] T. A. Courtade, "A strong entropy power inequality," *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 2173–2192, 2017.

[22] S. G. Bobkov, "Upper bounds for Fisher information," *Electronic Journal of Probability*, vol. 27, pp. 1–44, 2022.