Forward and reverse entropy power inequalities

Jiange Li

Einstein Institute of Mathematics
Hebrew University

High Dimensional Probability
06/17/2020

Joint work with Arnaud Marsiglietti and James Melbourne

Basic notations

Let X be a random vector in $\mathbb{R}^d$ with density f with respect to the Lebesgue measure. For $p \in [0, \infty]$, the p-Rényi entropy of X is

$$h_p(X) = \frac{1}{1-p} \log \int_{\mathbb{R}^n} f(x)^p dx.$$

Let X be a random vector in $\mathbb{R}^d$ with density f with respect to the Lebesgue measure. For $p \in [0, \infty]$, the p-Rényi entropy of X is

$$h_p(X) = \frac{1}{1-p} \log \int_{\mathbb{R}^n} f(x)^p dx.$$

Remark: For $p \in \{0, 1, \infty\}$, the definition is understood in the limiting sense; that is, $h_0(X) = \log |\mathrm{supp}(f)|$, Shannon-Boltzmann entropy $h(X) := h_1(X) = -\int_{\mathbb{R}^d} f(x) \log f(x) dx$ and $h_\infty(X) = -\log \|f\|_\infty$. By Jensen's inequality, $h_p(X)$ is non-increasing with respect to p.

Shannon's entropy power inequality (EPI)

Shannon'48: Let X, Y be independent random vectors in $\mathbb{R}^d$ such that the entropies of X, Y and $X+Y$ exist. We have

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

- Various proofs include the use of de Bruijn identity, Fisher information, harmonic analysis, Brunn-Minkoski with restricted sum, MMSE, optimal transport.
- Importance is well known: bounding capacity of communication channels, functional inequalities in probability, convex geometry and additive combinatorics, as well as in physics.

EPI for p-Rényi entropy: $p > 1$

Bobkov-Chistyakov'15: For any $p > 1$, there exists c_p such that for any independent random vectors $X_1, \cdots, X_n$ in $\mathbb{R}^d$,

$$e^{\frac{2}{d} h_p(X_1+\cdots+X_n)} \geq c_p \cdot \left(e^{\frac{2}{d} h_p(X_1)} + \cdots + e^{\frac{2}{d} h_p(X_n)}\right).$$

Remark: One can take $c_p = p^{\frac{1}{p-1}}/e$, which decreases from 1 to $1/e$, sharpened by Ram-Sason'16, Madiman-Melbourne-Xu'17.

EPI for p-Rényi entropy: $p > 1$, cont'd

Bobkov-Marsiglietti'17, L.'18: Let X, Y be independent random vectors in $\mathbb{R}^n$ such that X, Y and $X+Y$ have finite p-Rényi entropies for $p > 1$. We have

$$e^{\frac{\alpha_p}{d} h(X+Y)} \geq e^{\frac{\alpha_p}{d} h(X)} + e^{\frac{\alpha_p}{d} h(Y)},$$

$$\alpha_p = 2 \left[1 + \frac{1}{\log 2} \left(\frac{p+1}{p-1} \log \frac{p+1}{2p} + \frac{1}{p-1} \log p \right) \right]^{-1}.$$

EPI for p-Rényi entropy: $p>1$, cont'd

Bobkov-Marsiglietti'17, L.'18: Let X, Y be independent random vectors in $\mathbb{R}^n$ such that X, Y and $X+Y$ have finite p-Rényi entropies for $p>1$. We have

$$e^{\frac{\alpha_p}{d}h(X+Y)} \geq e^{\frac{\alpha_p}{d}h(X)} + e^{\frac{\alpha_p}{d}h(Y)},$$

$$\alpha_p = 2\left[1+\frac{1}{\log 2}\left(\frac{p+1}{p-1}\log\frac{p+1}{2p}+\frac{1}{p-1}\log p\right)\right]^{-1}.$$

Remark: It is unknown whether this inequality is sharp. The function α_p is monotone increasing. As $p \to 1$, we have $\alpha_p \to 2$ and, for large p, we have $\alpha_p \approx \frac{2(p-1)}{\log_2 p}$, which is asymptotically optimal up to a multiplicative constant.

The proof sketch

Both types of EPIs for Rényi entropy can be proved in a manner similar to Lieb's proof of Shannon's EPI.

The proof sketch

Both types of EPIs for Rényi entropy can be proved in a manner similar to Lieb's proof of Shannon's EPI. Firstly, we establish their equivalent linearized forms.

The proof sketch

Both types of EPIs for Rényi entropy can be proved in a manner similar to Lieb's proof of Shannon's EPI. Firstly, we establish their equivalent linearized forms. Secondly, we use the monotonicity of $h_p(X)$ to rewrite sharp Young's convolution inequality in an information-theoretic way, which is comparable with the linearized Rényi EPIs.

The proof sketch

Both types of EPIs for Rényi entropy can be proved in a manner similar to Lieb's proof of Shannon's EPI. Firstly, we establish their equivalent linearized forms. Secondly, we use the monotonicity of $h_p(X)$ to rewrite sharp Young's convolution inequality in an information-theoretic way, which is comparable with the linearized Rényi EPIs. Then, the multiplicative constant c_p or the entropy power α_p can be written as the maximum/minimum of some explicit functions.

The proof sketch

Both types of EPIs for Rényi entropy can be proved in a manner similar to Lieb's proof of Shannon's EPI. Firstly, we establish their equivalent linearized forms. Secondly, we use the monotonicity of $h_p(X)$ to rewrite sharp Young's convolution inequality in an information-theoretic way, which is comparable with the linearized Rényi EPIs. Then, the multiplicative constant c_p or the entropy power α_p can be written as the maximum/minimum of some explicit functions.

Remark: For $0 < p < 1$, the monotonicity of $h_p(X)$ does not work in the right direction. This technical issue prohibits us from proving EPIs for Rényi entropy of order $0 < p < 1$.

EPI for p-Rényi entropy: $0 < p < 1$

L.-Marsiglietti-Melbourne'19: For any $0 < p < 1$ and $\varepsilon > 0$, there exist i.i.d. random vectors $X_1, \cdots, X_n$ in $\mathbb{R}^d$ for some $d \geq 1$ and n large enough, such that

$$e^{\frac{2}{d} h_p(X_1 + \cdots + X_n)} < \varepsilon \cdot \left(e^{\frac{2}{d} h_p(X_1)} + \cdots + e^{\frac{2}{d} h_p(X_n)} \right).$$

EPI for p-Rényi entropy: $0 < p < 1$

L.-Marsiglietti-Melbourne'19: For any $0 < p < 1$ and $\varepsilon > 0$, there exist i.i.d. random vectors $X_1, \cdots, X_n$ in $\mathbb{R}^d$ for some $d \geq 1$ and n large enough, such that

$$e^{\frac{2}{d} h_p(X_1 + \cdots + X_n)} < \varepsilon \cdot \left(e^{\frac{2}{d} h_p(X_1)} + \cdots + e^{\frac{2}{d} h_p(X_n)} \right).$$

Remark: Such random vectors are essentially truncations of some random vector with finite covariance matrix, but infinite p-Rényi entropy for $0 < p < 1$. It is crucial that the p-Rényi entropy of their normalized sum converges to the p-Rényi entropy of the standard Gaussian. After normalization, the LHS is finite, but the RHS can be as large as possible.

EPI for p-Rényi entropy: $0 < p < 1$, cont'd

L.-Marsiglietti-Melbourne'19: For any $s \in (-1/d, 0)$ and $p \in (-sd, 1)$, there exists $c := c(s,p,d,n)$ such that for all independent random vectors $X_1, \cdots, X_n$ in $\mathbb{R}^d$ with s-concave densities,

$$e^{\frac{2}{d}h_p(X_1+\cdots+X_n)} \geq c \cdot \left(e^{\frac{2}{d}h_p(X_1)} + \cdots + e^{\frac{2}{d}h_p(X_n)}\right).$$

L.-Marsiglietti-Melbourne'19: For any $s \in (-1/d, 0)$, there exist $p_0 := p_0(s,d) \in (0,1)$ and $\alpha := \alpha(s,p,d)$ such that for $p \in (p_0, 1)$ and independent random vectors X, Y in $\mathbb{R}^d$ with s-concave densities,

$$e^{\frac{\alpha}{d}h_p(X+Y)} \geq e^{\frac{\alpha}{d}h_p(X)} + e^{\frac{\alpha}{d}h_p(Y)}.$$

Bobkov-Madiman'12: For independent log-concave random vectors X, Y in $\mathbb{R}^d$, there exist linear, volume preserving maps T_1, T_2 and an absolute constant c such that

$$e^{\frac{2}{d}h(T_1(X)+T_2(Y))} \leq c \cdot \left(e^{\frac{2}{d}h(X)} + e^{\frac{2}{d}h(Y)}\right).$$

Remark: This is a functional analogue of the well-known reverse Brunn-Minkowski inequality by V. Milman'86. The selection of T_1, T_2 is related to the slicing problem.

Busemann's Theorem: a reverse ∞-Rényi EPI

Busemann'49: Let $K \subset \mathbb{R}^d$ be a symmetric convex body. The radial function

$$\rho(v) = |K \cap v^\perp|, \; v \in \mathbb{S}^{d-1}$$

defines a symmetric convex body $I(K)$, which is called the intersection body of K.

Busemann'49: Let $K \subset \mathbb{R}^d$ be a symmetric convex body. The radial function

$$\rho(v) = |K \cap v^{\perp}|, \; v \in \mathbb{S}^{d-1}$$

defines a symmetric convex body $I(K)$, which is called the intersection body of K.

Entropy power version: Let X be a random vector uniformly distributed in K. For any $u, v \in \mathbb{R}^d$,

$$e^{h_\infty(\langle u+v, X\rangle)} \leq e^{h_\infty(\langle u, X\rangle)} + e^{h_\infty(\langle v, X\rangle)}.$$

Extensions to log-concave random vectors, Ball'88.

Conjecture (Gardner-Giannopoulos'99): Let $K \subset \mathbb{R}^d$ be a symmetric convex body. For any $p > -1$, the radial function

$$\rho(v) = \left(\int_K |K \cap (x + v^\perp)|^p dx \right)^{1/p}, \ v \in \mathbb{S}^{d-1}$$

defines a symmetric convex body $C_p(K)$, which is called the p-cross-section body of K.

Remark: Busemann's Theorem corresponds to the case $p = \infty$.

A reverse p-Rényi EPI?

Conjecture (Entropy power version 1): Let X be a random vector uniformly distributed in K. Let $p > -1$. For any $u, v \in \mathbb{R}^d$,

$$e^{h_{1+p}(\langle u+v, X\rangle)} \leq e^{h_{1+p}(\langle u, X\rangle)} + e^{h_{1+p}(\langle v, X\rangle)}.$$

A reverse p-Rényi EPI?

Conjecture (Entropy power version 1): Let X be a random vector uniformly distributed in K. Let $p > -1$. For any $u, v \in \mathbb{R}^d$,

$$e^{h_{1+p}(\langle u+v, X\rangle)} \le e^{h_{1+p}(\langle u, X\rangle)} + e^{h_{1+p}(\langle v, X\rangle)}.$$

Conjecture (Entropy power version 2): Let $(X, Y) \in \mathbb{R}^2$ be a symmetric log-concave random vector. For any $p > -1$, we have

$$e^{h_{1+p}(X+Y)} \le e^{h_{1+p}(X)} + e^{h_{1+p}(Y)}.$$

A reverse p-Rényi EPI?

Conjecture (Entropy power version 1): Let X be a random vector uniformly distributed in K. Let $p > -1$. For any $u, v \in \mathbb{R}^d$,

$$e^{h_{1+p}(\langle u+v, X\rangle)} \le e^{h_{1+p}(\langle u, X\rangle)} + e^{h_{1+p}(\langle v, X\rangle)}.$$

Conjecture (Entropy power version 2): Let $(X, Y) \in \mathbb{R}^2$ be a symmetric log-concave random vector. For any $p > -1$, we have

$$e^{h_{1+p}(X+Y)} \le e^{h_{1+p}(X)} + e^{h_{1+p}(Y)}.$$

Remark: No assumption on independence between X and Y. The Shannon case was independently asked by Ball-Nayar-Tkocz'16.

Towards the conjecture

Ball-Nayar-Tkocz'16: For any symmetric log-concave random vector $(X, Y) \in \mathbb{R}^2$, we have

$$e^{\frac{1}{5}h(X+Y)} \leq e^{\frac{1}{5}h(X)} + e^{\frac{1}{5}h(Y)}.$$

Towards the conjecture

Ball-Nayar-Tkocz'16: For any symmetric log-concave random vector $(X, Y) \in \mathbb{R}^2$, we have

$$e^{\frac{1}{5}h(X+Y)} \leq e^{\frac{1}{5}h(X)} + e^{\frac{1}{5}h(Y)}.$$

L'18: For any log-concave random vector $(X, Y) \in \mathbb{R}^2$, we have

$$e^{h_2(X+Y)} \leq e^{h_2(X)} + e^{h_2(Y)}.$$

Towards the conjecture

Ball-Nayar-Tkocz'16: For any symmetric log-concave random vector $(X, Y) \in \mathbb{R}^2$, we have

$$e^{\frac{1}{5}h(X+Y)} \leq e^{\frac{1}{5}h(X)} + e^{\frac{1}{5}h(Y)}.$$

L'18: For any log-concave random vector $(X, Y) \in \mathbb{R}^2$, we have

$$e^{h_2(X+Y)} \leq e^{h_2(X)} + e^{h_2(Y)}.$$

Remark: Symmetry is not required. The proof crucially relies on the simple observation that $h_2(X) = h_\infty(X - Y)$, where Y is an independent copy of X. The exponent in the Shannon case can be improved from $1/5$ to $1/3$.

A geometric functional inequality?

The Shannon case of the conjecture is implied by the following geometric functional inequality

$$\int (\varphi_1')^2 f - \left(\int \varphi_1' f \right)^2 \leq \int \left(\frac{(f')^2}{f} - f'' \right) (\varphi_2 - \varphi_1^2),$$

where f is the density of X and

$$\varphi_1(x) = \mathbb{E}[Y|X = x], \quad \varphi_2(x) = \mathbb{E}[Y^2|X = x].$$

Thanks for your attention!