Entropy Power Inequalities: Results and Speculation

Venkat Anantharam

EECS Department
University of California, Berkeley

December 12, 2013
Workshop on Coding and Information Theory
Institute of Mathematical Research
The University of Hong Kong

Contributions involve joint work with Varun Jog

Outline

1. Setting the stage
2. The Brunn Minkowski inequality
3. Proofs of the EPI
4. Mrs. Gerber's Lemma
5. Young's inequality
6. Versions of the EPI for discrete random variables
7. EPI for Groups or Order 2^n
8. The Fourier transform on a finite abelian group
9. Young's inequality on a finite abelian group
10. Brunn Minkowski inequality on a finite abelian group
11. Speculation

Outline

1. Setting the stage
2. The Brunn Minkowski inequality
3. Proofs of the EPI
4. Mrs. Gerber's Lemma
5. Young's inequality
6. Versions of the EPI for discrete random variables
7. EPI for Groups or Order 2^n
8. The Fourier transform on a finite abelian group
9. Young's inequality on a finite abelian group
10. Brunn Minkowski inequality on a finite abelian group
11. Speculation

Differential entropy and entropy power

Differential entropy and entropy power

- Let $\mathbf{X}$ be an $\mathbb{R}^n$-valued random variable with density f.

Differential entropy and entropy power

- Let $\mathbf{X}$ be an $\mathbb{R}^n$-valued random variable with density f.
- The *differential entropy* of $\mathbf{X}$ is

$$h(\mathbf{X}) := E[-\log f(\mathbf{X})] \ .$$

Differential entropy and entropy power

- Let $\mathbf{X}$ be an $\mathbb{R}^n$-valued random variable with density f.
- The *differential entropy* of $\mathbf{X}$ is

$$h(\mathbf{X}) := E[-\log f(\mathbf{X})] \ .$$

- The *entropy power* of $\mathbf{X}$ is

$$N(\mathbf{X}) := \frac{1}{2\pi e} e^{\frac{2}{n} h(\mathbf{X})} \ .$$

Shannon's entropy power inequality

Shannon's entropy power inequality

Theorem (Entropy Power Inequality)

If $\mathbf{X}$ *and* $\mathbf{Y}$ *are independent* $\mathbb{R}^n$*-valued random variables, then*

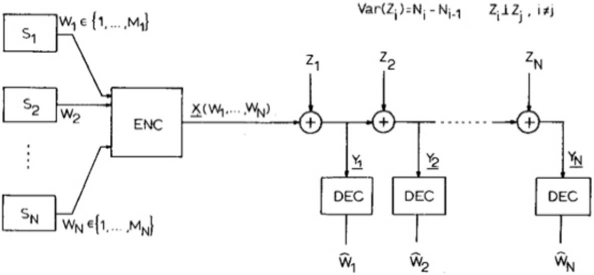
$$e^{\frac{2}{n}h(\mathbf{X}+\mathbf{Y})} \geq e^{\frac{2}{n}h(\mathbf{X})} + e^{\frac{2}{n}h(\mathbf{Y})},$$

with equality if and only if $\mathbf{X}$ *and* $\mathbf{Y}$ *are Gaussian with proportional covariance matrices.*

Notable applications of the EPI-1

A Simple Converse for Broadcast Channels with Additive White Gaussian Noise

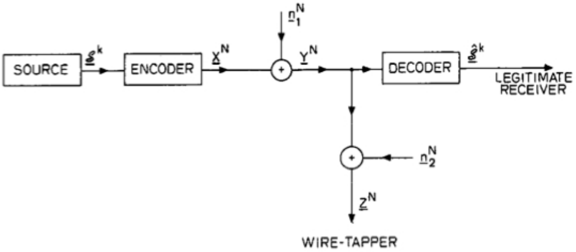
PATRICK P. BERGMANS



Notable applications of the EPI-2

The Gaussian Wire-Tap Channel

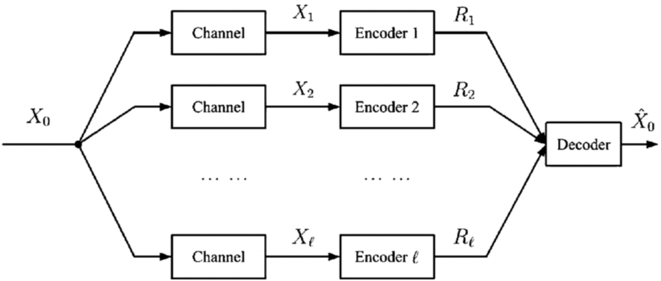
S. K. LEUNG-YAN-CHEONG, MEMBER, IEEE, AND MARTIN E. HELLMAN, MEMBER, IEEE



Notable applications of the EPI-3

The Rate-Distortion Function for the Quadratic Gaussian CEO Problem

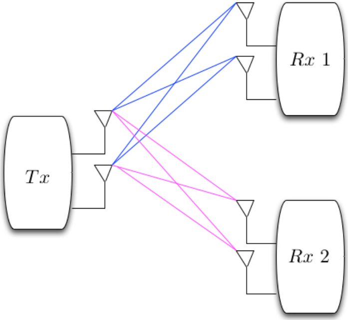
Yasutada Oohama



Notable applications of the EPI-4

The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel

Hanan Weingarten, *Student Member, IEEE*, Yossef Steinberg, *Member, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

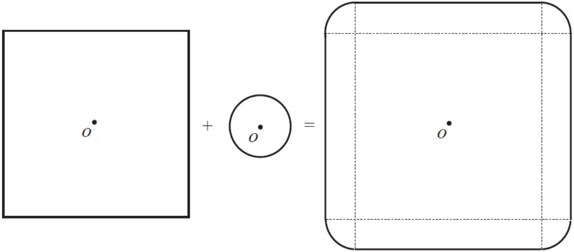


Outline

1. Setting the stage
2. The Brunn Minkowski inequality
3. Proofs of the EPI
4. Mrs. Gerber's Lemma
5. Young's inequality
6. Versions of the EPI for discrete random variables
7. EPI for Groups or Order 2^n
8. The Fourier transform on a finite abelian group
9. Young's inequality on a finite abelian group
10. Brunn Minkowski inequality on a finite abelian group
11. Speculation

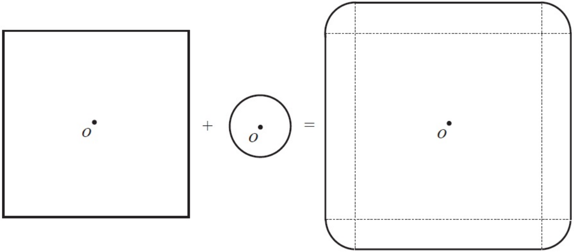
The Minkowski sum of two subsets of $\mathbb{R}^n$

The Minkowski sum of two subsets of $\mathbb{R}^n$



Minkowski sum of a square and a circle

The Minkowski sum of two subsets of $\mathbb{R}^n$



Minkowski sum of a square and a circle

- In general the Minkowski sum of two sets $A, B \subseteq \mathbb{R}^n$ is

$$A + B := \{a + b \ : \ a \in A \text{ and } b \in B\} \ .$$

The Brunn Minkowski inequality

The Brunn Minkowski inequality

- Let $A, B \subset \mathbb{R}^n$ be convex bodies (compact convex sets with nonempty interior).

The Brunn Minkowski inequality

- Let $A, B \subset \mathbb{R}^n$ be convex bodies (compact convex sets with nonempty interior).
- Then

$$\mathrm{Vol}(A+B)^{\frac{1}{n}} \geq \mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}} \ .$$

The Brunn Minkowski inequality

- Let $A, B \subset \mathbb{R}^n$ be convex bodies (compact convex sets with nonempty interior).
- Then
$$\mathrm{Vol}(A+B)^{\frac{1}{n}} \geq \mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}} \ .$$
- Equality holds iff A and B are homothetic, i.e. equal up to translation and dilation.

The Brunn Minkowski inequality

- Let $A, B \subset \mathbb{R}^n$ be convex bodies (compact convex sets with nonempty interior).
- Then

$$\mathrm{Vol}(A+B)^{\frac{1}{n}} \geq \mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}} \ .$$

- Equality holds iff A and B are homothetic, i.e. equal up to translation and dilation.
- An equivalent form is that for convex bodies $A, B \subset \mathbb{R}^n$ and $0 < \lambda < 1$ we have

$$\mathrm{Vol}((1-\lambda)A + \lambda B)^{\frac{1}{n}} \geq (1-\lambda)\mathrm{Vol}(A)^{\frac{1}{n}} + \lambda \mathrm{Vol}(B)^{\frac{1}{n}} \ .$$

The Brunn Minkowski inequality

- Let $A, B \subset \mathbb{R}^n$ be convex bodies (compact convex sets with nonempty interior).
- Then
$$\text{Vol}(A+B)^{\frac{1}{n}} \geq \text{Vol}(A)^{\frac{1}{n}} + \text{Vol}(B)^{\frac{1}{n}} .$$
- Equality holds iff A and B are homothetic, i.e. equal up to translation and dilation.
- An equivalent form is that for convex bodies $A, B \subset \mathbb{R}^n$ and $0 < \lambda < 1$ we have
$$\text{Vol}((1-\lambda)A + \lambda B)^{\frac{1}{n}} \geq (1-\lambda)\text{Vol}(A)^{\frac{1}{n}} + \lambda \text{Vol}(B)^{\frac{1}{n}} .$$
- Another equivalent form is that for convex bodies $A, B \subset \mathbb{R}^n$ and $0 < \lambda < 1$ we have
$$\text{Vol}((1-\lambda)A + \lambda B)^{\frac{1}{n}} \geq \min(\text{Vol}(A), \text{Vol}(B)) .$$

Parallels between the EPI and the Brunn Minkowski inequality

Parallels between the EPI and the Brunn Minkowski inequality

- The distribution of a random variable corresponds to a (measurable) subset of $\mathbb{R}^n$.

Parallels between the EPI and the Brunn Minkowski inequality

- The distribution of a random variable corresponds to a (measurable) subset of $\mathbb{R}^n$.
- Addition of independent random variables corresponds to the Minkowski sum.

Parallels between the EPI and the Brunn Minkowski inequality

- The distribution of a random variable corresponds to a (measurable) subset of $\mathbb{R}^n$.
- Addition of independent random variables corresponds to the Minkowski sum.
- Balls play the role of spherically symmetric Gaussians.

Parallels between the EPI and the Brunn Minkowski inequality

- The distribution of a random variable corresponds to a (measurable) subset of $\mathbb{R}^n$.
- Addition of independent random variables corresponds to the Minkowski sum.
- Balls play the role of spherically symmetric Gaussians.
- For instance, an equivalent form of the EPI is that if $\mathbf{X}$ and $\mathbf{Y}$ are independent $\mathbb{R}^n$-valued random variables, and $0 < \lambda < 1$, then

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

Outline

1 Setting the stage
2 The Brunn Minkowski inequality
3 Proofs of the EPI
4 Mrs. Gerber's Lemma
5 Young's inequality
6 Versions of the EPI for discrete random variables
7 EPI for Groups or Order 2^n
8 The Fourier transform on a finite abelian group
9 Young's inequality on a finite abelian group
10 Brunn Minkowski inequality on a finite abelian group
11 Speculation

Some history

Some history

- The EPI was conjectured by Shannon, who showed that $\mathbf{X}$ and $\mathbf{Y}$ being Gaussian with proportional covariance matrices is a stationary point of the difference between the entropy power of the sum and sum of the entropy powers. This did not exclude the possibility of it being a local maximum or a saddle point.

Some history

- The EPI was conjectured by Shannon, who showed that **X** and **Y** being Gaussian with proportional covariance matrices is a stationary point of the difference between the entropy power of the sum and sum of the entropy powers. This did not exclude the possibility of it being a local maximum or a saddle point.
- Stam gave the first rigorous proof of the EPI in 1959, using de Bruijn's identity which relates Fisher information and differential entropy.

Some history

- The EPI was conjectured by Shannon, who showed that $\mathbf{X}$ and $\mathbf{Y}$ being Gaussian with proportional covariance matrices is a stationary point of the difference between the entropy power of the sum and sum of the entropy powers. This did not exclude the possibility of it being a local maximum or a saddle point.
- Stam gave the first rigorous proof of the EPI in 1959, using de Bruijn's identity which relates Fisher information and differential entropy.
- Blachman's 1965 paper gives a simplified and succinct version of Stam's proof.

Some history

- The EPI was conjectured by Shannon, who showed that $\mathbf{X}$ and $\mathbf{Y}$ being Gaussian with proportional covariance matrices is a stationary point of the difference between the entropy power of the sum and sum of the entropy powers. This did not exclude the possibility of it being a local maximum or a saddle point.
- Stam gave the first rigorous proof of the EPI in 1959, using de Bruijn's identity which relates Fisher information and differential entropy.
- Blachman's 1965 paper gives a simplified and succinct version of Stam's proof.
- Subsequently, several proofs have appeared, some of which will be mentioned here for our story.

Blachman's (1965) proof strategy for the EPI following Stam (1959)

Blachman's (1965) proof strategy for the EPI following Stam (1959)

Step 1: Prove de Bruijn's identity for scalar random variables

Blachman's (1965) proof strategy for the EPI following Stam (1959)

Step 1: Prove de Bruijn's identity for scalar random variables

- Let $X_t = X + \sqrt{t}N$, where $N \sim \mathcal{N}(0,1)$. Let $X \sim p(x)$, and $X + \sqrt{t}N \sim p_t(x_t)$.

Blachman's (1965) proof strategy for the EPI following Stam (1959)

Step 1: Prove de Bruijn's identity for scalar random variables

- Let $X_t = X + \sqrt{t}N$, where $N \sim \mathcal{N}(0,1)$. Let $X \sim p(x)$, and $X + \sqrt{t}N \sim p_t(x_t)$.

De Bruijn's Identity

$$\frac{d}{dt}h(X_t) = \frac{1}{2}J(X_t) = \frac{1}{2}\int_{-\infty}^{\infty}\left(\frac{\partial}{\partial x_t}p_t(x_t)\right)^2 \frac{dx_t}{p_t(x_t)}$$

Blachman's (1965) proof strategy for the EPI following Stam (1959)

Step 1: Prove de Bruijn's identity for scalar random variables

- Let $X_t = X + \sqrt{t}N$, where $N \sim \mathcal{N}(0,1)$. Let $X \sim p(x)$, and $X + \sqrt{t}N \sim p_t(x_t)$.

De Bruijn's Identity

$$\frac{d}{dt}h(X_t) = \frac{1}{2}J(X_t) = \frac{1}{2}\int_{-\infty}^{\infty}(\frac{\partial}{\partial x_t}p_t(x_t))^2\frac{dx_t}{p_t(x_t)}$$

- The proof follows from differentiating

$$h(X_t) = -\int_{-\infty}^{\infty} p_t(x_t)\log p_t(x_t)dx_t,$$

with respect to t, and using the heat equation

$$\frac{\partial}{\partial t}p_t(x_t) = \frac{1}{2}\frac{\partial^2}{\partial t^2}p_t(x_t).$$

Blachman's proof strategy (contd.)

Blachman's proof strategy (contd.)

Step 2: Prove Stam's inequality

Blachman's proof strategy (contd.)

Step 2: Prove Stam's inequality

- If X and Y are independent, and $Z = X + Y$ then

Stam's Inequality

$\frac{1}{J(Z)} \geq \frac{1}{J(X)} + \frac{1}{J(Y)}$

Blachman's proof strategy (contd.)

Step 2: Prove Stam's inequality

- If X and Y are independent, and $Z = X + Y$ then

Stam's Inequality

$\frac{1}{J(Z)} \geq \frac{1}{J(X)} + \frac{1}{J(Y)}$

- The proof goes by showing that

$$\mathbb{E}\left(a\frac{p'_X(x)}{p_X(x)} + b\frac{p'_Y(y)}{p_Y(y)}|Z = z\right) = (a+b)\frac{p'_Z(z)}{p_Z(z)},$$

for all a, b. Then squaring both sides, using Jensen's inequality and optimizing over the choice of a, b establishes the inequality.

Blachman's proof strategy (contd.)

Blachman's proof strategy (contd.)

Step 3: Set up the one parameter flow defined by the heat equation starting with the two initial distributions, with suitable time reparametrizations $f(t), g(t)$ for each, respectively

Blachman's proof strategy (contd.)

Step 3: Set up the one parameter flow defined by the heat equation starting with the two initial distributions, with suitable time reparametrizations $f(t), g(t)$ for each, respectively

- Let $f(t)$ and $g(t)$ be increasing functions tending to infinity with t, and consider the function

$$s(t) = \frac{e^{2h(X_f)} + e^{2h(Y_g)}}{e^{2h(Z_{f+g})}}.$$

Blachman's proof strategy (contd.)

Step 3: Set up the one parameter flow defined by the heat equation starting with the two initial distributions, with suitable time reparametrizations $f(t), g(t)$ for each, respectively

- Let $f(t)$ and $g(t)$ be increasing functions tending to infinity with t, and consider the function

$$s(t) = \frac{e^{2h(X_f)} + e^{2h(Y_g)}}{e^{2h(Z_{f+g})}}.$$

- Intuitively, $\lim_{t\to\infty} s(t) = 1$, since both initial distributions become increasingly Gaussian as time progresses along the flow.

Blachman's proof strategy (contd.)

Step 3: Set up the one parameter flow defined by the heat equation starting with the two initial distributions, with suitable time reparametrizations $f(t), g(t)$ for each, respectively

- Let $f(t)$ and $g(t)$ be increasing functions tending to infinity with t, and consider the function

$$s(t) = \frac{e^{2h(X_f)} + e^{2h(Y_g)}}{e^{2h(Z_{f+g})}}.$$

- Intuitively, $\lim_{t\to\infty} s(t) = 1$, since both initial distributions become increasingly Gaussian as time progresses along the flow.
- Choosing $f'(t) = exp(2h(X_f))$ and $g'(t) = exp(2h(Y_g))$, and using Stam's inequality gives $s'(t) \geq 0$. This gives $s(0) \leq 1$, proving the EPI.

Lieb's EPI proof as interpreted by Verdú and Guo

Lieb's EPI proof as interpreted by Verdú and Guo

- Reparametrization of the de Bruijn identity gives the Guo-Shamai-Verdú I-MMSE relation:

$$\frac{d}{d\gamma} I(X; \sqrt{\gamma} X + N) = \frac{1}{2} \mathrm{mmse}(X, \gamma),$$

and thus

$$h(X) = \frac{1}{2} \log 2\pi e - \frac{1}{2} \int_0^\infty \frac{1}{1+\gamma} - \mathrm{mmse}(X, \gamma) d\gamma.$$

Lieb's EPI proof as interpreted by Verdú and Guo

- Reparametrization of the de Bruijn identity gives the Guo-Shamai-Verdú I-MMSE relation:

$$\frac{d}{d\gamma} I(X; \sqrt{\gamma} X + N) = \frac{1}{2} \mathrm{mmse}(X, \gamma),$$

  and thus

$$h(X) = \frac{1}{2} \log 2\pi e - \frac{1}{2} \int_0^\infty \frac{1}{1+\gamma} - \mathrm{mmse}(X, \gamma) d\gamma.$$

- The EPI in its equivalent form,

$$h(X_1 \cos\alpha + X_2 \sin\alpha) \geq \cos^2\alpha\, h(X_1) + \sin^2\alpha\, h(X_2),$$

  can be proved using these ideas.

Verdú and Guo's proof: Proof strategy

Verdú and Guo's proof: Proof strategy

$$h(X_1 \cos\alpha + X_2 \sin\alpha) \geq \cos^2\alpha h(X_1) + \sin^2\alpha h(X_2) =$$

$$\frac{1}{2}\int \Big(\text{mmse}(X_1 \cos\alpha + X_2 \sin\alpha) - \cos^2\alpha \ \text{mmse}(X_1, \gamma) + \sin^2\alpha \ \text{mmse}(X_2, \gamma)\Big) d\gamma$$

Verdú and Guo's proof: Proof strategy

$$h(X_1 \cos\alpha + X_2 \sin\alpha) \geq \cos^2\alpha h(X_1) + \sin^2\alpha h(X_2) =$$
$$\frac{1}{2}\int \Big(\text{mmse}(X_1 \cos\alpha + X_2 \sin\alpha) - \cos^2\alpha \;\text{mmse}(X_1, \gamma)$$
$$+ \sin^2\alpha \;\text{mmse}(X_2, \gamma)\Big) d\gamma$$

- Consider $Z_1 = \sqrt{\gamma}X_1 + N_1$, $Z_2 = \sqrt{\gamma}X_2 + N_2$ and $Z = Z_1 \cos\alpha + Z_2 \sin\alpha$. Then

$$\text{mmse}(X_1 \cos\alpha + X_2 \sin\alpha | Z) \geq \text{mmse}(X | Z_1, Z_2)$$

immediately gives the term inside the integral is ≥ 0.

Szarek and Voiculescu's proof of the EPI

Szarek and Voiculescu's proof of the EPI

- Szarek and Voiculescu (1996) give a proof of the EPI working directly with typical sets, using a 'restricted' version of the Brunn-Minkowski inequality.

Szarek and Voiculescu's proof of the EPI

- Szarek and Voiculescu (1996) give a proof of the EPI working directly with typical sets, using a 'restricted' version of the Brunn-Minkowski inequality.
- The restricted Minkowski sum of sets A and B, based on $\Theta \subseteq A \times B$ is defined as

$$A +_{\Theta} B := \{x + y : (x, y) \in \Theta\} .$$

Szarek and Voiculescu's proof of the EPI

- Szarek and Voiculescu (1996) give a proof of the EPI working directly with typical sets, using a 'restricted' version of the Brunn-Minkowski inequality.
- The restricted Minkowski sum of sets A and B, based on $\Theta \subseteq A \times B$ is defined as

$$A +_\Theta B := \{x + y : (x, y) \in \Theta\} \ .$$

- Szarek and Voiculescu's restricted B-M inequality says: For every ϵ, there exists δ such that if Θ is large enough, viz $V_{2n}(\Theta) \geq (1-\delta)^n V_n(A) V_n(B)$, then

$$V_n(A +_\Theta B)^{\frac{2}{n}} \geq (1-\epsilon)(V_n(A)^{\frac{2}{n}} + V_n(B)^{\frac{2}{n}}).$$

Szarek and Voiculescu's proof (contd.)

Szarek and Voiculescu's proof (contd.)

- To prove the EPI, Szarek and Voiculescu replace A and B by typical sets for n i.i.d. copies of X and Y respectively, and define Θ as all pairs (x^n, y^n) (each marginal typical) such that $x^n + y^n$ is typical for $X + Y$.

Szarek and Voiculescu's proof (contd.)

- To prove the EPI, Szarek and Voiculescu replace A and B by typical sets for n i.i.d. copies of X and Y respectively, and define Θ as all pairs (x^n, y^n) (each marginal typical) such that $x^n + y^n$ is typical for $X + Y$.
- For this choice of restriction, they determine suitable sequence of typicality defining constants (ϵ_n, δ_n) going to 0, thus proving the EPI.

Outline

1 Setting the stage
2 The Brunn Minkowski inequality
3 Proofs of the EPI
4 Mrs. Gerber's Lemma
5 Young's inequality
6 Versions of the EPI for discrete random variables
7 EPI for Groups or Order 2^n
8 The Fourier transform on a finite abelian group
9 Young's inequality on a finite abelian group
10 Brunn Minkowski inequality on a finite abelian group
11 Speculation

Mrs. Gerber's Lemma for Gaussian random variables

Mrs. Gerber's Lemma for Gaussian random variables

- $\frac{n}{2}\log(e^{\frac{2}{n}x} + e^{\frac{2}{n}y})$ is the minimum achievable differential entropy of $\mathbf{X} + \mathbf{Y}$ where $\mathbf{X}$ and $\mathbf{Y}$ are independent $\mathbb{R}^n$-valued random variables with differential entropies x and y respectively.

Mrs. Gerber's Lemma for Gaussian random variables

- $\frac{n}{2}\log(e^{\frac{2}{n}x} + e^{\frac{2}{n}y})$ is the minimum achievable differential entropy of $\mathbf{X} + \mathbf{Y}$ where $\mathbf{X}$ and $\mathbf{Y}$ are independent $\mathbb{R}^n$-valued random variables with differential entropies x and y respectively.
- Note that
$$(x, y) \mapsto \frac{n}{2}\log(e^{\frac{2}{n}x} + e^{\frac{2}{n}y})$$
is convex in (x, y).

Mrs. Gerber's Lemma for Gaussian random variables

- $\frac{n}{2}\log(e^{\frac{2}{n}x} + e^{\frac{2}{n}y})$ is the minimum achievable differential entropy of $\mathbf{X} + \mathbf{Y}$ where $\mathbf{X}$ and $\mathbf{Y}$ are independent $\mathbb{R}^n$-valued random variables with differential entropies x and y respectively.
- Note that
$$(x, y) \mapsto \frac{n}{2}\log(e^{\frac{2}{n}x} + e^{\frac{2}{n}y})$$
is convex in (x, y).
- In particular, for fixed x,
$$y \mapsto \frac{n}{2}\log(e^{\frac{2}{n}x} + e^{\frac{2}{n}y})$$
is convex in y and for fixed y,
$$x \mapsto \frac{n}{2}\log(e^{\frac{2}{n}x} + e^{\frac{2}{n}y})$$
is convex in x.

Outline

1. Setting the stage
2. The Brunn Minkowski inequality
3. Proofs of the EPI
4. Mrs. Gerber's Lemma
5. Young's inequality
6. Versions of the EPI for discrete random variables
7. EPI for Groups or Order 2^n
8. The Fourier transform on a finite abelian group
9. Young's inequality on a finite abelian group
10. Brunn Minkowski inequality on a finite abelian group
11. Speculation

Young's inequality

Young's inequality

- The *Hausdorff-Young inequality* gives an estimate on the norm of the Fourier transform: $\|\mathcal{F}f\|_{p'} \leq \|f\|_p \,, \quad \text{for } 1 \leq p \leq 2.$

Young's inequality

- The *Hausdorff-Young inequality* gives an estimate on the norm of the Fourier transform: $\|\mathcal{F}f\|_{p'} \le \|f\|_p \ , \quad$ for $1 \le p \le 2$.
- The sharp constant in the inequality was found by Beckner

$$\|\mathcal{F}f\|_{p'} \le C_p \|f\|_p \ , \quad \text{where } C_p := \sqrt{\frac{|p|^{1/p}}{|p'|^{1/p'}}} \ .$$

Young's inequality

- The *Hausdorff-Young inequality* gives an estimate on the norm of the Fourier transform: $\|\mathcal{F}f\|_{p'} \le \|f\|_p$, for $1 \le p \le 2$.
- The sharp constant in the inequality was found by Beckner

$$\|\mathcal{F}f\|_{p'} \le C_p \|f\|_p \ , \quad \text{where } C_p := \sqrt{\frac{|p|^{1/p}}{|p'|^{1/p'}}} \ .$$

- This leads to *Young's inequality*. If $p, q, r > 0$ satisfy $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, then

$$\|f * g\|_r \le \frac{C_p C_q}{C_r} \|f\|_p \|g\|_q \ , \quad \text{if } p, q, r \ge 1 \ ,$$

$$\|f * g\|_r \ge \frac{C_p C_q}{C_r} \|f\|_p \|g\|_q \ , \quad \text{if } 1 \ge p, q, r \ .$$

Young's inequality

- The *Hausdorff-Young inequality* gives an estimate on the norm of the Fourier transform: $\|\mathcal{F}f\|_{p'} \leq \|f\|_p$, for $1 \leq p \leq 2$.
- The sharp constant in the inequality was found by Beckner

$$\|\mathcal{F}f\|_{p'} \leq C_p \|f\|_p \ , \quad \text{where } C_p := \sqrt{\frac{|p|^{1/p}}{|p'|^{1/p'}}} \ .$$

- This leads to *Young's inequality*. If $p, q, r > 0$ satisfy $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, then

$$\|f * g\|_r \leq \frac{C_p C_q}{C_r} \|f\|_p \|g\|_q \ , \quad \text{if } p, q, r \geq 1 \ ,$$

$$\|f * g\|_r \geq \frac{C_p C_q}{C_r} \|f\|_p \|g\|_q \ , \quad \text{if } 1 \geq p, q, r \ .$$

- The second half of this is called the *reverse* Young's inequality.

EPI from the Young's inequality

EPI from the Young's inequality

- For $p > 1$, $h_p(\mathbf{X}) := \frac{p}{1-p}\|f\|_p$ is called the *Renyi entropy* of $\mathbf{X} \sim f$.

EPI from the Young's inequality

- For $p > 1$, $h_p(\mathbf{X}) := \frac{p}{1-p}\|f\|_p$ is called the *Renyi entropy* of $\mathbf{X} \sim f$.
- $N_p(\mathbf{X}) := \frac{1}{2\pi} p^{-p'/p} \|f\|_p^{-2p'/n}$ is the *Renyi entropy power of* $\mathbf{X} \sim f$.

EPI from the Young's inequality

- For $p > 1$, $h_p(\mathbf{X}) := \frac{p}{1-p}\|f\|_p$ is called the *Renyi entropy* of $\mathbf{X} \sim f$.
- $N_p(\mathbf{X}) := \frac{1}{2\pi} p^{-p'/p}\|f\|_p^{-2p'/n}$ is the *Renyi entropy power of* $\mathbf{X} \sim f$.
- As $p \to 1$ these converge respectively to the differential entropy and the entropy power of $\mathbf{X}$.

EPI from the Young's inequality

- For $p > 1$, $h_p(\mathbf{X}) := \frac{p}{1-p} \|f\|_p$ is called the *Renyi entropy* of $\mathbf{X} \sim f$.
- $N_p(\mathbf{X}) := \frac{1}{2\pi} p^{-p'/p} \|f\|_p^{-2p'/n}$ is the *Renyi entropy power of* $\mathbf{X} \sim f$.
- As $p \to 1$ these converge respectively to the differential entropy and the entropy power of $\mathbf{X}$.
- Given $r > 1$ and $0 < \lambda < 1$ define $p(r) := \frac{r}{(1-\lambda)+\lambda r}$ and $q(r) := \frac{r}{\lambda+(1-\lambda) r}$.

EPI from the Young's inequality

- For $p > 1$, $h_p(\mathbf{X}) := \frac{p}{1-p} \|f\|_p$ is called the *Renyi entropy* of $\mathbf{X} \sim f$.
- $N_p(\mathbf{X}) := \frac{1}{2\pi} p^{-p'/p} \|f\|_p^{-2p'/n}$ is the *Renyi entropy power of* $\mathbf{X} \sim f$.
- As $p \to 1$ these converge respectively to the differential entropy and the entropy power of $\mathbf{X}$.
- Given $r > 1$ and $0 < \lambda < 1$ define $p(r) := \frac{r}{(1-\lambda)+\lambda r}$ and $q(r) := \frac{r}{\lambda+(1-\lambda)r}$.
- Young's inequality gives

$$N_r(\mathbf{X}+\mathbf{Y}) \geq \left(\frac{N_p(\mathbf{X})}{1-\lambda}\right)^{1-\lambda} \left(\frac{N_q(\mathbf{Y})}{\lambda}\right)^{\lambda} .$$

EPI from the Young's inequality

- For $p > 1$, $h_p(\mathbf{X}) := \frac{p}{1-p}\|f\|_p$ is called the *Renyi entropy* of $\mathbf{X} \sim f$.
- $N_p(\mathbf{X}) := \frac{1}{2\pi} p^{-p'/p} \|f\|_p^{-2p'/n}$ is the *Renyi entropy power of* $\mathbf{X} \sim f$.
- As $p \to 1$ these converge respectively to the differential entropy and the entropy power of $\mathbf{X}$.
- Given $r > 1$ and $0 < \lambda < 1$ define $p(r) := \frac{r}{(1-\lambda)+\lambda r}$ and $q(r) := \frac{r}{\lambda+(1-\lambda)r}$.
- Young's inequality gives

$$N_r(\mathbf{X}+\mathbf{Y}) \geq \left(\frac{N_p(\mathbf{X})}{1-\lambda}\right)^{1-\lambda} \left(\frac{N_q(\mathbf{Y})}{\lambda}\right)^{\lambda} .$$

- Optimize over λ and let $r \to 1$ to get the EPI.

Brunn Minkowski inequality from the reverse Young's inequality

Brunn Minkowski inequality from the reverse Young's inequality

- The limit as $r \to 0$ in Young's inequality leads to the *Prékopa-Leindler inequality*: If $0 < \lambda < 1$ and f, g, h are nonnegative integrable functions on $\mathbb{R}^n$ satisfying

$$h((1-\lambda)x + \lambda y) \geq f(x)^{1-\lambda} g(y)^{\lambda}$$

for all $x, y \in \mathbb{R}^n$, then

$$\int_{\mathbb{R}^n} h(x)dx \geq \left(\int_{\mathbb{R}^n} f(x)dx \right)^{1-\lambda} \left(\int_{\mathbb{R}^n} g(x)dx \right)^{\lambda} .$$

Brunn Minkowski inequality from the reverse Young's inequality

- The limit as $r \to 0$ in Young's inequality leads to the *Prékopa-Leindler inequality*: If $0 < \lambda < 1$ and f, g, h are nonnegative integrable functions on $\mathbb{R}^n$ satisfying

$$h((1-\lambda)x + \lambda y) \geq f(x)^{1-\lambda} g(y)^{\lambda}$$

for all $x, y \in \mathbb{R}^n$, then

$$\int_{\mathbb{R}^n} h(x)dx \geq \left(\int_{\mathbb{R}^n} f(x)dx \right)^{1-\lambda} \left(\int_{\mathbb{R}^n} g(x)dx \right)^{\lambda} .$$

- Setting $f := 1_A$, $g := 1_B$, and $h := 1_{(1-\lambda)A+\lambda B}$, this gives

$$\mathrm{Vol}((1-\lambda)A+\lambda B) \geq \mathrm{Vol}(A)^{1-\lambda}\mathrm{Vol}(B)^{\lambda} \geq \min(\mathrm{Vol}(A), \mathrm{Vol}(B)) .$$

Outline

1 Setting the stage

2 The Brunn Minkowski inequality

3 Proofs of the EPI

4 Mrs. Gerber's Lemma

5 Young's inequality

6 Versions of the EPI for discrete random variables

7 EPI for Groups or Order 2^n

8 The Fourier transform on a finite abelian group

9 Young's inequality on a finite abelian group

10 Brunn Minkowski inequality on a finite abelian group

11 Speculation

Mrs. Gerber's lemma

Mrs. Gerber's lemma

- Let h be the binary entropy function, and h^{-1} be its inverse. Wyner and Ziv (1973) showed that for a binary X and arbitrary U, if $Z \sim Bern(p)$ is independent of (X, U) then

$$h(X \oplus Z|U) \geq h(h^{-1}(H(X|U) \star p).$$

Mrs. Gerber's lemma

- Let h be the binary entropy function, and h^{-1} be its inverse. Wyner and Ziv (1973) showed that for a binary X and arbitrary U, if $Z \sim Bern(p)$ is independent of (X, U) then

$$h(X \oplus Z|U) \geq h(h^{-1}(H(X|U) \star p).$$

- The result follows immediately from the following lemma, using Jensen's inequality -

Lemma

The function

$$x \mapsto h(h^{-1}(x) \star p)$$

is convex on $0 \leq x \leq \log 2$.

Wyner and Ziv's proof of Mrs. Gerber's lemma

Wyner and Ziv's proof of Mrs. Gerber's lemma

Introduce the parametrization $\alpha = \frac{1-h^{-1}(x)}{2}$, let $p = \frac{1-a}{2}$, and differentiate wrt α, to get that

$$f''(x) \geq 0 \iff a(1-\alpha^2)\log\frac{1+\alpha}{1-\alpha} \leq (1-a\alpha)^2\log\frac{1+a\alpha}{1-a\alpha},$$

which can be verified using the series expansion of $\log\frac{1+x}{1-x}$.

Shamai and Wyner's EPI

Shamai and Wyner's EPI

- Shamai and Wyner show that -

EPI for binary sequences

For *binary independent processes* with entropies $H(X)$ and $H(Y)$,

$$\sigma(Z) \geq \sigma(X) \star \sigma(Y),$$

where $Z = X \oplus Y$ and $\sigma(X) = h^{-1}(H(X)), \sigma(Y) = h^{-1}(H(Y)$ and $\sigma(Z) = h^{-1}(H(Z))$.

Shamai and Wyner's proof strategy for the binary EPI

Shamai and Wyner's proof strategy for the binary EPI

$$
\begin{aligned}
H(Z_0|Z_{-n}^{-1}) \geq H(Z_0|Z_{-n}^{-1}, X_{-n}^{-1}, Y_{-n}^{-1}) &= H(Z_0|X_{-n}^{-1}, Y_{-n}^{-1}), \\
&= \sum P(X_{-n}^{-1} = x)P(Y_{-n}^{-1} = y)H(Z_0|X_{-n}^{-1} = x, Y_{-n}^{-1} = y) \\
&= \sum P(X_{-n}^{-1} = x)P(Y_{-n}^{-1} = y)h(\alpha(x) \star \beta(y))
\end{aligned}
$$

Where $\alpha(x) = P(X_0 = 1|X_{-n}^{-1} = x)$, $\beta(x) = P(Y_0 = 1|Y_{-n}^{-1} = y)$.

Shamai and Wyner's proof strategy for the binary EPI

$$\begin{aligned}
H(Z_0|Z_{-n}^{-1}) \geq H(Z_0|Z_{-n}^{-1}, X_{-n}^{-1}, Y_{-n}^{-1}) &= H(Z_0|X_{-n}^{-1}, Y_{-n}^{-1}), \\
&= \sum P(X_{-n}^{-1} = x)P(Y_{-n}^{-1} = y)H(Z_0|X_{-n}^{-1} = x, Y_{-n}^{-1} = y) \\
&= \sum P(X_{-n}^{-1} = x)P(Y_{-n}^{-1} = y)h(\alpha(x) \star \beta(y))
\end{aligned}$$

Where $\alpha(x) = P(X_0 = 1|X_{-n}^{-1} = x)$, $\beta(x) = P(Y_0 = 1|Y_{-n}^{-1} = y)$.

- Using convexity in MGL twice, summing over x and then y, we get

$$H(Z_0|Z_{-n}^{-1}) \geq h(h^{-1}(H(X_0|X_{-n}^{-1}) \star h^{-1}(H(Y_0|Y_{-n}^{-1})).$$

Taking $n \to \infty$ proves the result.

Harremoës's EPI for the Binomial family

Harremoës's EPI for the Binomial family

- Let $X_n \sim Binomial(n, \frac{1}{2})$, then for all $m, n \geq 1$,

$$e^{2H(X_n+X_m)} \geq e^{2H(X_n)} + e^{2H(X_m)}.$$

Harremoës's EPI for the Binomial family

- Let $X_n \sim Binomial(n, \frac{1}{2})$, then for all $m, n \geq 1$,

$$e^{2H(X_n+X_m)} \geq e^{2H(X_n)} + e^{2H(X_m)}.$$

- The proof follows by showing that $Y_n = \frac{e^{2H(X_n)}}{n}$ is an increasing sequence, and thus is super additive, i.e $Y_{m+n} \geq Y_m + Y_n$.

n	1	2	3	4	5	6
$\frac{e^{2H(X_n)}}{n}$	4	4	4.105	4.173	4.212	4.233

Figure : Values of Y_n

Our approach

Our approach

Suppose the random variables take values on a finite abelian group G. We define the function $f_G : [0, \log |G|] \times [0, \log |G|] \to [0, \log |G|]$ by

$$f_G(x, y) = \min_{H(X)=x, H(Y)=y} H(X + Y)$$

Our approach

Suppose the random variables take values on a finite abelian group G. We define the function $f_G : [0, \log |G|] \times [0, \log |G|] \to [0, \log |G|]$ by

$$f_G(x, y) = \min_{H(X)=x, H(Y)=y} H(X + Y)$$

We now ask the questions-

- Does f_G have a succinct description?

Our approach

Suppose the random variables take values on a finite abelian group G.
We define the function $f_G : [0, \log|G|] \times [0, \log|G|] \to [0, \log|G|]$ by

$$f_G(x, y) = \min_{H(X)=x, H(Y)=y} H(X + Y)$$

We now ask the questions-

- Does f_G have a succinct description?
- Does f_G have any nice properties?

Our approach

Suppose the random variables take values on a finite abelian group G.
We define the function $f_G : [0, \log|G|] \times [0, \log|G|] \to [0, \log|G|]$ by

$$f_G(x, y) = \min_{H(X)=x, H(Y)=y} H(X + Y)$$

We now ask the questions-

- Does f_G have a succinct description?
- Does f_G have any nice properties?

We try to exploit the group structure to answer these questions.

Mimicking Blachman's proof?

Mimicking Blachman's proof?

- Find a 1-parameter group to evolve the probability distributions towards the uniform distribution on the group?

Mimicking Blachman's proof?

- Find a 1-parameter group to evolve the probability distributions towards the uniform distribution on the group?
- Perhaps the analogs of the Gaussian distributions are the distributions of the form

$$p(0) = 1 - \epsilon \ , \quad p(g) = \frac{1}{|G| - 1}\epsilon \text{ for } g \neq 0 \ , \quad 0 \leq \epsilon \leq \frac{|G| - 1}{|G|} \ ?$$

Mimicking Blachman's proof?

- Find a 1-parameter group to evolve the probability distributions towards the uniform distribution on the group?
- Perhaps the analogs of the Gaussian distributions are the distributions of the form

$$p(0) = 1 - \epsilon \ , \quad p(g) = \frac{1}{|G|-1}\epsilon \text{ for } g \neq 0 \ , \quad 0 \leq \epsilon \leq \frac{|G|-1}{|G|} \ ?$$

- But these are not extremal for the EPI. For instance, for $G = \mathbb{Z}_2 \oplus \mathbb{Z}_2$, if ϵ is chosen so that $h(\epsilon) + \epsilon \log 3 = \log 2$ (note that $\epsilon < \frac{3}{4}$), then we have

$$h(1 - 2\epsilon + \frac{4}{3}\epsilon^2) + 2\epsilon(1 - \frac{2}{3}\epsilon) \log 3 > \log 2 \ .$$

Mimicking Blachman's proof?

- Find a 1-parameter group to evolve the probability distributions towards the uniform distribution on the group?
- Perhaps the analogs of the Gaussian distributions are the distributions of the form

$$p(0) = 1 - \epsilon \ , \quad p(g) = \frac{1}{|G| - 1} \epsilon \text{ for } g \neq 0 \ , \quad 0 \leq \epsilon \leq \frac{|G| - 1}{|G|} \ ?$$

- But these are not extremal for the EPI. For instance, for $G = \mathbb{Z}_2 \oplus \mathbb{Z}_2$, if ϵ is chosen so that $h(\epsilon) + \epsilon \log 3 = \log 2$ (note that $\epsilon < \frac{3}{4}$), then we have

$$h(1 - 2\epsilon + \frac{4}{3}\epsilon^2) + 2\epsilon(1 - \frac{2}{3}\epsilon) \log 3 > \log 2 \ .$$

- However, for the uniform distribution on a subgroup, its convolution with itself has entropy $\log 2$.

Mimicking Lieb's proof?

Mimicking Lieb's proof?

- Given group valued random variables X_1 and X_2, what is the analog of $X_1 \cos \alpha + X_2 \sin \alpha$?

Mimicking Lieb's proof?

- Given group valued random variables X_1 and X_2, what is the analog of $X_1 \cos\alpha + X_2 \sin\alpha$?
- There is a version of Stam's inequality for finite abelian groups (Gibilisco and Isola, 2008). Given a set of generators $\Gamma := \{\gamma_1, \gamma_2, \ldots, \gamma_m\}$ for the group G, define the "Fisher information" of a random variable X by

$$J(X) := \sum_{\gamma \in \Gamma} \sum_{g \in G} \left(\frac{p_X(g) - p_X(\gamma^{-1} g)}{p_X(g)} \right)^2 p_X(g) \ .$$

Then one has $\frac{1}{J(X_1+X_2)} \geq \frac{1}{J(X_1)} + \frac{1}{J(X_2)}$ for independent G-valued random variables X_1 and X_2.

Mimicking Lieb's proof?

- Given group valued random variables X_1 and X_2, what is the analog of $X_1 \cos\alpha + X_2 \sin\alpha$?
- There is a version of Stam's inequality for finite abelian groups (Gibilisco and Isola, 2008). Given a set of generators $\Gamma := \{\gamma_1, \gamma_2, \dots, \gamma_m\}$ for the group G, define the "Fisher information" of a random variable X by

$$J(X) := \sum_{\gamma \in \Gamma} \sum_{g \in G} \left(\frac{p_X(g) - p_X(\gamma^{-1} g)}{p_X(g)} \right)^2 p_X(g) \ .$$

  Then one has $\frac{1}{J(X_1+X_2)} \geq \frac{1}{J(X_1)} + \frac{1}{J(X_2)}$ for independent G-valued random variables X_1 and X_2.
- However this notion of "Fisher information" is merely mimicking formal properties of the Fisher information from the continuous case. It seems to have little to do with estimation, which, for finite groups ought to refer to likelihood ratio type quantities.

Mimicking Szarek and Voiculescu's proof?

Mimicking Szarek and Voiculescu's proof?

- This is a promising direction.

Mimicking Szarek and Voiculescu's proof?

- This is a promising direction.
- The question remains, what is the correct analog of the restricted Minkowski sum?

Mimicking Szarek and Voiculescu's proof?

- This is a promising direction.
- The question remains, what is the correct analog of the restricted Minkowski sum?
- The key lemma driving the proof of Szarek and Voiculescu appears to be Euclidean in character.

Outline

1 Setting the stage
2 The Brunn Minkowski inequality
3 Proofs of the EPI
4 Mrs. Gerber's Lemma
5 Young's inequality
6 Versions of the EPI for discrete random variables
7 EPI for Groups or Order 2^n
8 The Fourier transform on a finite abelian group
9 Young's inequality on a finite abelian group
10 Brunn Minkowski inequality on a finite abelian group
11 Speculation

Our results

Our results

- For a group G of order 2^n, we explicitly describe f_G in terms of $f_{\mathbb{Z}_2}$.

Our results

- For a group G of order 2^n, we explicitly describe f_G in terms of $f_{\mathbb{Z}_2}$.
- We also describe the minimum entropy achieving distributions on G, these distributions are analogs of Gaussians in this sense.

Our results

- For a group G of order 2^n, we explicitly describe f_G in terms of $f_{\mathbb{Z}_2}$.
- We also describe the minimum entropy achieving distributions on G, these distributions are analogs of Gaussians in this sense.
- The $f_G(x, y)$ function obtained has the property that it is convex in x for a fixed y. This is yet another version of Mrs. Gerber's Lemma.

Statement of our results

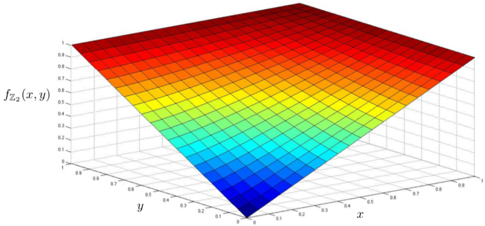
Statement of our results

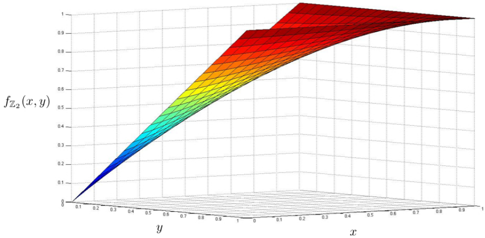
Theorem

f_G depends only on the size of G, and is denoted by f_{2^n}, where

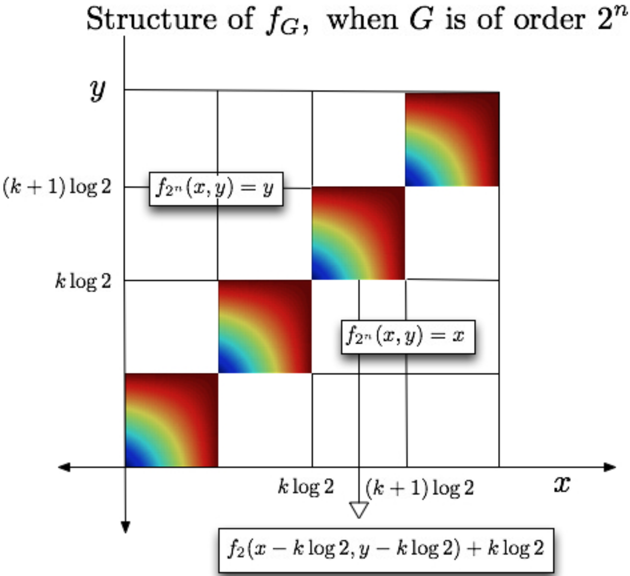
$$f_{2^n}(x,y) = \begin{cases} f_2(x - k\log 2, y - k\log 2) + k\log 2, \\ \qquad \text{if } k\log 2 \le x, y \le (k+1)\log 2 \ , \\ \max(x,y) \quad \text{otherwise.} \end{cases}$$

Visualizing our results





Visualizing our results (contd.)



A conjecture

A conjecture

Generalized Mrs. Gerber's Lemma

If G is a finite group, then $f_G(x, y)$ is convex in x for a fixed y, and by symmetry convex in y for a fixed x.

A conjecture

Generalized Mrs. Gerber's Lemma

If G is a finite group, then $f_G(x, y)$ is convex in x for a fixed y, and by symmetry convex in y for a fixed x.

- If one is less optimistic, one might make this conjecture only for finite abelian groups.

A conjecture

Generalized Mrs. Gerber's Lemma

If G is a finite group, then $f_G(x, y)$ is convex in x for a fixed y, and by symmetry convex in y for a fixed x.

- If one is less optimistic, one might make this conjecture only for finite abelian groups.
- Simulations for $\mathbb{Z}_3$ and $\mathbb{Z}_5$ seem to support this conjecture.

A conjecture

Generalized Mrs. Gerber's Lemma

If G is a finite group, then $f_G(x, y)$ is convex in x for a fixed y, and by symmetry convex in y for a fixed x.

- If one is less optimistic, one might make this conjecture only for finite abelian groups.
- Simulations for $\mathbb{Z}_3$ and $\mathbb{Z}_5$ seem to support this conjecture.
- We saw already that f_{2^n} satisfies the conjecture.

EPI for abelian groups of order 2^n: proof strategy

EPI for abelian groups of order 2^n: proof strategy

- First observe that for $\mathbb{Z}_2$ the function $f_{\mathbb{Z}_2}(x, y)$ is explicitly given by
$$f_{\mathbb{Z}_2}(x, y) = h(h^{-1}(x) \star h^{-1}(y)).$$

EPI for abelian groups of order 2^n: proof strategy

- First observe that for $\mathbb{Z}_2$ the function $f_{\mathbb{Z}_2}(x, y)$ is explicitly given by
$$f_{\mathbb{Z}_2}(x, y) = h(h^{-1}(x) \star h^{-1}(y)).$$

- The proof proceeds by first proving the EPI for $\mathbb{Z}_4$, using convexity properties of $f_{\mathbb{Z}_2}$ and concavity of entropy.

EPI for abelian groups of order 2^n: proof strategy

- First observe that for $\mathbb{Z}_2$ the function $f_{\mathbb{Z}_2}(x, y)$ is explicitly given by

$$f_{\mathbb{Z}_2}(x, y) = h(h^{-1}(x) \star h^{-1}(y)).$$

- The proof proceeds by first proving the EPI for $\mathbb{Z}_4$, using convexity properties of $f_{\mathbb{Z}_2}$ and concavity of entropy.
- We then use induction, where we assume that the result holds for all groups of size 2^n and prove it groups of size 2^{n+1}.

EPI for abelian groups of order 2^n: proof strategy

- First observe that for $\mathbb{Z}_2$ the function $f_{\mathbb{Z}_2}(x, y)$ is explicitly given by

$$f_{\mathbb{Z}_2}(x, y) = h(h^{-1}(x) \star h^{-1}(y)).$$

- The proof proceeds by first proving the EPI for $\mathbb{Z}_4$, using convexity properties of $f_{\mathbb{Z}_2}$ and concavity of entropy.
- We then use induction, where we assume that the result holds for all groups of size 2^n and prove it groups of size 2^{n+1}.
- The induction part of the proof has almost exactly the same structure as the proof for $\mathbb{Z}_4$. Thus proving the result for $\mathbb{Z}_4$ is the key step in our proof.

Proof for $\mathbb{Z}_4$

Proof for $\mathbb{Z}_4$

- For $\mathbb{Z}_4$, splitting the support of distributions on $\mathbb{Z}_4$ into two parts, $\{0, 2\}$ (even part) and $\{1, 3\}$ (odd part) and using precisely two ingredients: concavity of entropy, and convexity in MGL we arrive at the lower bound

Lower bound

$$f_4(x, y) \geq \min_{u,v} f_2(u, v) + f_2(x - u, y - v)$$

Proof for $\mathbb{Z}_4$

- For $\mathbb{Z}_4$, splitting the support of distributions on $\mathbb{Z}_4$ into two parts, $\{0,2\}$ (even part) and$\{1,3\}$ (odd part) and using precisely two ingredients: concavity of entropy, and convexity in MGL we arrive at the lower bound

Lower bound

$$f_4(x,y) \geq \min_{u,v} f_2(u,v) + f_2(x-u,y-v)$$

- To evaluate the minimum, we prove some additional properties of $f_{\mathbb{Z}_2}$, specifically regarding its behavior along lines through the origin. The key Lemma we use is

Lemma

$\frac{\partial f_{\mathbb{Z}_2}}{\partial x}$ (and by symmetry, $\frac{\partial f_{\mathbb{Z}_2}}{\partial y}$) strictly decreases along lines through the origin

Proving the lemma

Proving the lemma

- We use the parametrization $h^{-1}(x) = p$ and $h^{-1}(y) = q$. Our strategy involves brute force differentiation, and a series of steps where we conclude that proving the lemma is equivalent to proving

$$
\begin{aligned}
F(p) = p^2(\log(p))^2 - (1-p)^2(\log(1-p))^2 + \\
(1-2p)\left(\log(p)\log(1-p) + p\log(p) + (1-p)\log(1-p)\right) \\
\leq 0
\end{aligned}
$$

Proving the lemma

- We use the parametrization $h^{-1}(x) = p$ and $h^{-1}(y) = q$. Our strategy involves brute force differentiation, and a series of steps where we conclude that proving the lemma is equivalent to proving

$$
\begin{aligned}
F(p) = p^2(\log(p))^2 - (1-p)^2(\log(1-p))^2 + \\
(1-2p)\left(\log(p)\log(1-p) + p\log(p) + (1-p)\log(1-p)\right) \\
\leq 0
\end{aligned}
$$

- We show that proving $F(p) \leq 0$ is the same as proving $\frac{d^5}{dp^5}F(p) \leq 0$, we differentiate F five times, use a polynomial approximation of log, and finally use Sturm's theorem to prove that the resulting polynomial is non-positive.

Outline

1. Setting the stage
2. The Brunn Minkowski inequality
3. Proofs of the EPI
4. Mrs. Gerber's Lemma
5. Young's inequality
6. Versions of the EPI for discrete random variables
7. EPI for Groups or Order 2^n
8. The Fourier transform on a finite abelian group
9. Young's inequality on a finite abelian group
10. Brunn Minkowski inequality on a finite abelian group
11. Speculation

Norm of the Fourier transform on a finite abelian group

Norm of the Fourier transform on a finite abelian group

- Every finite abelian group G has a dual group $\hat{G}$ comprised of the characters of G, i.e. the homomorphisms from G to the unit circle in the complex plane.

Norm of the Fourier transform on a finite abelian group

- Every finite abelian group G has a dual group $\hat{G}$ comprised of the characters of G, i.e. the homomorphisms from G to the unit circle in the complex plane.
- $\hat{G}$ has the same cardinality and structure as G and G can be thought of as the dual of $\hat{G}$.

Norm of the Fourier transform on a finite abelian group

- Every finite abelian group G has a dual group $\hat{G}$ comprised of the characters of G, i.e. the homomorphisms from G to the unit circle in the complex plane.
- $\hat{G}$ has the same cardinality and structure as G and G can be thought of as the dual of $\hat{G}$.
- The Fourier transform on the group is the map from functions on G to functions on $\hat{G}$ given by $\mathcal{F}(f)(\gamma) := \frac{1}{|G|^{-1/2}} \sum_g \gamma(g) f(g)$.

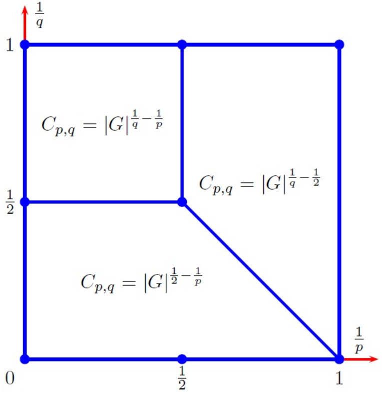
Norm of the Fourier transform on a finite abelian group

- Every finite abelian group G has a dual group $\hat{G}$ comprised of the characters of G, i.e. the homomorphisms from G to the unit circle in the complex plane.
- $\hat{G}$ has the same cardinality and structure as G and G can be thought of as the dual of $\hat{G}$.
- The Fourier transform on the group is the map from functions on G to functions on $\hat{G}$ given by $\mathcal{F}(f)(\gamma) := \frac{1}{|G|^{-1/2}} \sum_g \gamma(g) f(g)$.
- With counting measure on G and $\hat{G}$, we have, for $p, q > 0$, $\mathcal{F} \; : \; L^p(G) \mapsto L^q(\hat{G})$. What is the norm of this map?

Norm of the Fourier transform on a finite abelian group

- Every finite abelian group G has a dual group $\hat{G}$ comprised of the characters of G, i.e. the homomorphisms from G to the unit circle in the complex plane.
- $\hat{G}$ has the same cardinality and structure as G and G can be thought of as the dual of $\hat{G}$.
- The Fourier transform on the group is the map from functions on G to functions on $\hat{G}$ given by $\mathcal{F}(f)(\gamma) := \frac{1}{|G|^{-1/2}} \sum_g \gamma(g) f(g)$.
- With counting measure on G and $\hat{G}$, we have, for $p, q > 0$, $\mathcal{F} \ : \ L^p(G) \mapsto L^q(\hat{G})$. What is the norm of this map?
- Resolved by Gilbert and Rzeszotnik (2010).

Visualizing the norm of the FT on a finite abelian group



Norm of the Fourier Transform on a finite abelian group

Outline

1. Setting the stage
2. The Brunn Minkowski inequality
3. Proofs of the EPI
4. Mrs. Gerber's Lemma
5. Young's inequality
6. Versions of the EPI for discrete random variables
7. EPI for Groups or Order 2^n
8. The Fourier transform on a finite abelian group
9. Young's inequality on a finite abelian group
10. Brunn Minkowski inequality on a finite abelian group
11. Speculation

Young's inequality on a finite abelian group

Young's inequality on a finite abelian group

- There is a forward Young's inequality, which reads: For every $p, q, r \geq 1$ such that $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$ we have

$$\|f * g\|_r \leq \|f\|_p \|g\|_q \ .$$

Outline

1 Setting the stage
2 The Brunn Minkowski inequality
3 Proofs of the EPI
4 Mrs. Gerber's Lemma
5 Young's inequality
6 Versions of the EPI for discrete random variables
7 EPI for Groups or Order 2^n
8 The Fourier transform on a finite abelian group
9 Young's inequality on a finite abelian group
10 Brunn Minkowski inequality on a finite abelian group
11 Speculation

Brunn Minkowski inequality on a finite abelian group

Brunn Minkowski inequality on a finite abelian group

- For a finite abelian group G, consider the function

$$\mu_G(r,s) := \min\{|A+B| \ : \ A, B \subset G, |A| = r, |B| = s\} \ .$$

Brunn Minkowski inequality on a finite abelian group

- For a finite abelian group G, consider the function

$$\mu_G(r,s) := \min\{|A+B| \ : \ A,B \subset G, |A|=r, |B|=s\} \ .$$

- Then we have $\mu_G(r,s) = \min_{d \text{ divides } |G|} \left(\lceil \frac{r}{d} \rceil + \lceil \frac{s}{d} \rceil - 1 \right) d$.

Brunn Minkowski inequality on a finite abelian group

- For a finite abelian group G, consider the function

$$\mu_G(r,s) := \min\{|A+B| \ : \ A, B \subset G, |A| = r, |B| = s\} \ .$$

- Then we have $\mu_G(r,s) = \min_{d \text{ divides } |G|} \left(\lceil \frac{r}{d} \rceil + \lceil \frac{s}{d} \rceil - 1 \right) d$.
- In particular for G a finite abelian group of size $n = 2^k$, we have

$$\mu_G(r,s) = r \circ s$$

where $r \circ s$ is the *Hopf-Stiefel function* defined as the smallest positive integer m for which the polynomial $(x+y)^m$ falls in the ideal generated by x^r and y^s in the polynomial ring $\mathbb{F}_2[x,y]$.

Brunn Minkowski inequality on a finite abelian group

- For a finite abelian group G, consider the function

$$\mu_G(r, s) := \min\{|A + B| \ : \ A, B \subset G, |A| = r, |B| = s\} \ .$$

- Then we have $\mu_G(r, s) = \min_{d \text{ divides } |G|} \left(\lceil \frac{r}{d} \rceil + \lceil \frac{s}{d} \rceil - 1 \right) d$.
- In particular for G a finite abelian group of size $n = 2^k$, we have

$$\mu_G(r, s) = r \circ s$$

where $r \circ s$ is the *Hopf-Stiefel function* defined as the smallest positive integer m for which the polynomial $(x + y)^m$ falls in the ideal generated by x^r and y^s in the polynomial ring $\mathbb{F}_2[x, y]$.
- Involved history of partial results. See Eliahu and Kervaire (2007) for the history.

Outline

1. Setting the stage
2. The Brunn Minkowski inequality
3. Proofs of the EPI
4. Mrs. Gerber's Lemma
5. Young's inequality
6. Versions of the EPI for discrete random variables
7. EPI for Groups or Order 2^n
8. The Fourier transform on a finite abelian group
9. Young's inequality on a finite abelian group
10. Brunn Minkowski inequality on a finite abelian group
11. Speculation

Questions

Questions

- What form does the reverse Young's inequality take on finite abelian groups?

Questions

- What form does the reverse Young's inequality take on finite abelian groups?
- Less ambitiously, what form does the reverse Young's inequality take on finite groups of size 2^k?

Questions

- What form does the reverse Young's inequality take on finite abelian groups?
- Less ambitiously, what form does the reverse Young's inequality take on finite groups of size 2^k?
- Is the EPI we proved for finite abelian groups of size 2^k derivable as a limit from the Young's inequality for such groups?

Questions

- What form does the reverse Young's inequality take on finite abelian groups?
- Less ambitiously, what form does the reverse Young's inequality take on finite groups of size 2^k?
- Is the EPI we proved for finite abelian groups of size 2^k derivable as a limit from the Young's inequality for such groups?
- Is the Brunn Minkowski inequality for finite abelian groups of size 2^k (expressed via the Hopf-Stiefel function) derivable as a limit from the appropriate reverse Young's inequality for such groups?

Questions

- What form does the reverse Young's inequality take on finite abelian groups?
- Less ambitiously, what form does the reverse Young's inequality take on finite groups of size 2^k?
- Is the EPI we proved for finite abelian groups of size 2^k derivable as a limit from the Young's inequality for such groups?
- Is the Brunn Minkowski inequality for finite abelian groups of size 2^k (expressed via the Hopf-Stiefel function) derivable as a limit from the appropriate reverse Young's inequality for such groups?
- Is the general Brunn Minkowski inequality for finite abelian groups derivable as a limit from the appropriate reverse Young's inequality for such groups?

Questions

- What form does the reverse Young's inequality take on finite abelian groups?
- Less ambitiously, what form does the reverse Young's inequality take on finite groups of size 2^k?
- Is the EPI we proved for finite abelian groups of size 2^k derivable as a limit from the Young's inequality for such groups?
- Is the Brunn Minkowski inequality for finite abelian groups of size 2^k (expressed via the Hopf-Stiefel function) derivable as a limit from the appropriate reverse Young's inequality for such groups?
- Is the general Brunn Minkowski inequality for finite abelian groups derivable as a limit from the appropriate reverse Young's inequality for such groups?
- Will a limit of the Young's inequality for a general finite Abelian group give rise to an EPI for each such group?

The End