An information-theoretic approach to Kneser-Poulsen conjecture in discrete geometry

Dongbin Li
(joint work with Gautam Aishwarya, Irfan Alam, Sergii Myroshnychenko, and Oscar Zatarain-Vera)

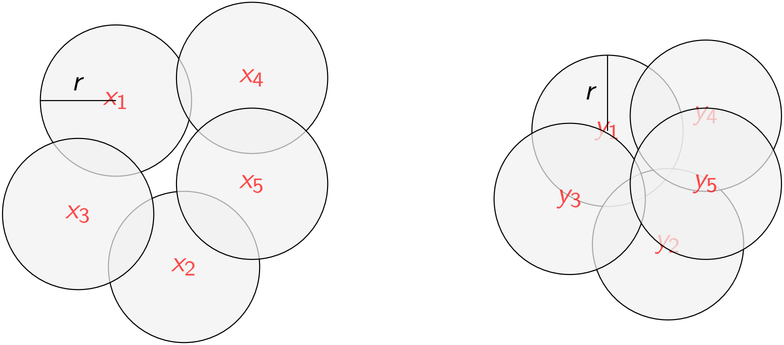
University of Delaware

December 13, 2022

Contents of the talk

- Background
- Our information-theoretic approach
- Some main results and sketch of proofs
- Open problems

Background



Conjecture (Poulsen 1954, Kneser 1955)

Let $K = \{x_1, \ldots, x_k\}$ *and* $L = \{y_1, \cdots, y_k\}$ *be two subsets in* $\mathbb{R}^d$ *such that* $\|x_i - x_j\|_2 \geq \|y_i - y_j\|_2$*, then the Kneser-Poulsen conjecture asserts that, for any* $r > 0$*,*

$$\mathrm{Vol}_d\left(\bigcup_{i=1}^{k} \mathcal{B}(x_i, r)\right) \geq \mathrm{Vol}_d\left(\bigcup_{i=1}^{k} \mathcal{B}(y_i, r)\right).$$

Conjecture (Gromov 1978, Klee-Wagon 1991)

Under the same assumptions,

$$\mathrm{Vol}_d\left(\bigcap_{i=1}^{k}\mathcal{B}(x_i, r)\right) \leq \mathrm{Vol}_d\left(\bigcap_{i=1}^{k}\mathcal{B}(y_i, r)\right).$$

Some Developments

1. The union conjecture holds for continuous contractions. Csikós(1998)

$$P(t) = \{p_1(t), \dots, p_k(t)\},\ 0 \le t \le 1$$

such that

$$P(0) = K, P(1) = L$$

$$K = \{x_1, \dots, x_k\} \subseteq \mathbb{R}^d$$
$$L = \{y_1, \dots, y_k\} \subseteq \mathbb{R}^d$$

and for all $1 \le i < j \le k$,

$$\|p_i(t) - p_j(t)\|_2 \text{ monotone decreasing in } t.$$

2. The union and intersection conjectures are true for $d = 2$, Bezdek-Connelly(2001).
3. Strong contractions and Uniform contractions, Bezdek-Naszódi(2018).

Survey paper: Four classic problems, Fejes Tóth-Kuperberg (2022).

Rephrase the Kneser-Poulsen Conjecture

$T : K \to \mathbb{R}^d$,

$$T(x_i) = y_i, \;\; i = 1, \ldots, k.$$

Then T is 1-Lip. (Recall $\|y_i - y_j\|_2 \le \|x_i - x_j\|_2$)

Definition (Minkowski sum)

$$A + B := \{a + b : a \in A, b \in B\}$$

$\mathcal{B}$ unit ball,

$$\bigcup_{i=1}^{k} \mathcal{B}(x_i, r) = K + r\mathcal{B}.$$

$$\bigcup_{i=1}^{k} \mathcal{B}(y_i, r) = T(K) + r\mathcal{B}.$$

$$\text{1-Lip } T : K \to \mathbb{R}^d \xRightarrow{\text{Kirszbraun's extension theorem}} \text{1-Lip } T : \mathbb{R}^d \to \mathbb{R}^d$$

Lip(T) ≤ 1

Conjecture

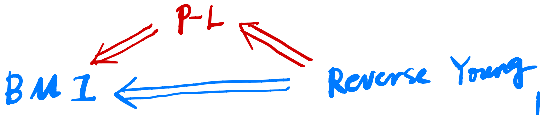
For every contraction T *of* $\mathbb{R}^d$ *and every compact set* $K \subseteq \mathbb{R}^d$, $r > 0$, *we have*

$$\mathrm{Vol}_d(T[K] + r\mathcal{B}) \leq \mathrm{Vol}_d(K + r\mathcal{B}).$$

Remark: It is unknown even for convex body K.

BMI: $\mathrm{Vol}_d(\lambda A + (1-\lambda)B) \geq \mathrm{Vol}_d^{\lambda}(A)\,\mathrm{Vol}_d^{1-\lambda}(B)$

$\lambda \in [0,1]$, A, B compact



$\|f * g\|_r \geq C^n \|f\|_p \|g\|_q$

$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$

$0 < p, q, r \leq 1$

(Young, 1912)

(Leindler, 1972)

(Beckner, 1975)

(Gardner: The BMI '02)

(Dembo-Cover-Thomas: Information theoretic inequalities, 1991)

Definition (Rényi entropies)

For a discrete r.v. X, the Rényi entropy of order $\alpha \in (0,1) \cup (1,\infty)$ is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_i p_i^\alpha \right) = \frac{\alpha}{1-\alpha} \log \|p\|_\alpha$$

For a $\mathbb{R}^d$-valued r.v.$X \sim f$, the Rényi entropy of order α is defined as

$$h_\alpha(f) = h_\alpha(X) := \frac{1}{1-\alpha} \log \int_{\mathbb{R}^d} f^\alpha(x) \mathrm{d}x.$$

Definition (Rényi entropies)

For a discrete r.v. X, the Rényi entropy of order $\alpha \in (0,1) \cup (1,\infty)$ is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_i p_i^\alpha \right).$$

For a $\mathbb{R}^d$-valued r.v.$X \sim f$, the Rényi entropy of order α is defined as

$$h_\alpha(f) = h_\alpha(X) := \frac{1}{1-\alpha} \log \int_{\mathbb{R}^d} f^\alpha(x) dx.$$

The Rényi entropies of order $\alpha = 0, 1, \infty$ are obtained by taking limits. In particular, let $\alpha \searrow 0$, we have

$$h_0(X) = \log \mathrm{Vol}_d(\mathrm{supp}(f)).$$

Useful fact: $h_\alpha(T(X)) = h_\alpha(X), \ T$ linear, $|\det(T)| = 1$.

Our information-theoretic approach

X, W, independent, $W \sim \text{Uniform}(\mathcal{B})$

$$\begin{cases} \text{Supp}(X) = K \\ \text{Supp}(W) = \mathcal{B} \end{cases} \implies \begin{cases} Supp(X + W) = K + \mathcal{B} \\ Supp(T(X) + W) = T(K) + \mathcal{B} \end{cases}$$

Our information-theoretic approach

X, W, independent, $W \sim \text{Uniform}(\mathcal{B})$

$$\begin{cases} \text{Supp}(X) = K \\ \text{Supp}(W) = \mathcal{B} \end{cases} \implies \begin{cases} Supp(X + W) = K + \mathcal{B} \\ Supp(T(X) + W) = T(K) + \mathcal{B} \end{cases}$$

For any fixed $\epsilon > 0$, $\forall \alpha \in (0, \epsilon)$

$$h_\alpha(T[X] + W) \le h_\alpha(X + W)$$

$$\Downarrow$$

$$\text{Vol}_d(T[K] + \mathcal{B}) \le \text{Vol}_d(K + \mathcal{B}).$$

Question (A broad information-theoretic question)

Let X and W be two independent $\mathbb{R}^d$-valued random variables. Further assume that W is log-concave and satisfies some symmetry property, say, radial symmetry. For a contraction $T : \mathbb{R}^d \to \mathbb{R}^d$, and $\alpha \in [0, \infty]$, under what additional assumptions do we have

$$h_\alpha(T(X) + W) \leq h_\alpha(X + W)\,?$$

Remark: An affirmative answer to this question at $\alpha = 0$ for any X supported on a finite set will prove the Kneser-Poulsen Conjecture.

What we have done

- Use of majorization to obtain Rényi entropic inequalities for all orders under convexity conditions of various flavors.

(Liu, Liu, Poor, Shamai '10)

- Exploiting a vector generalization of Costa's entropy power inequality (EPI) and variance comparisons to obtain results in a stronger form in some cases when W is Gaussian and $\alpha = 1$.

- Use of metric distortion of entropy to give a clear proof for $\alpha = 2$ in full generality.

Definition (Log-concavity)

A function $f : \mathbb{R}^d \mapsto [0, \infty)$ is log-concave if f can be written as $f(x) = e^{-U(x)}$, where $U : \mathbb{R}^d \mapsto (-\infty, +\infty]$ is a convex function. We call a $\mathbb{R}^d$ valued random variable X log-concave if its density is log-concave.

Examples:

$$f(x) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{\|x\|_2^2}{2}}, \quad f(x) = \frac{\mathbb{1}_K(x)}{|K|}, \ K \text{ convex body.}$$

Theorem (Prékopa)

If f and g are two nonnegative integrable log-concave functions on $\mathbb{R}^d$, then $f \star g$ is also log-concave.

Strategy

Example:
$K = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d, X_1,\ X_2$ two random variables.

$$X_1 : \mathbb{P}\{X = x_i\} = 1 \text{ for some } i,$$

$$X_2 : X_2 \sim \text{Uniform}(K).$$

For $\forall \alpha \in [0, \infty]$,

$$H_\alpha(X_1) = 0, \quad H_\alpha(X_2) = \log n.$$

More "peaked" distribution $\Longrightarrow$ Smaller Rényi entropies.

$$\Delta_n = \{p \in \mathbb{R}^n : \sum_{i=1}^{n} p_i = 1,\ p_i \geq 0, i = 1, \ldots, n.\}$$

For $p \in \Delta_n$,

$$\tilde{p},\ \text{same components, but sorted in descending order.}$$

For $p, q \in \Delta_n$, p is said to be majorized by q, written as $p \prec q$, iff

$$\sum_{i=1}^{k} \tilde{p}_i \leq \sum_{i=1}^{k} \tilde{q}_i, k = 1, \ldots, n.$$

And

$$p \prec q \Rightarrow H_\alpha(p) \geq H_\alpha(q), \forall \alpha \in [0, \infty].$$

Definition (Symmetric Decreasing Rearrangement)

Let A be a measurable set of finite volume in $\mathbb{R}^d$. Its symmetric rearrangement A^* is the open ball centered at the origin whose volume agrees with A. Define the SDR of the indicator function as follows:

$$\mathbb{1}_A^*(x) = \mathbb{1}_{A^*}(x).$$

Let f be a non-negative integrable function. The symmetric decreasing rearrangement f^* of f is defined as,

$$f^*(x) = \int_0^\infty \mathbb{1}^*_{\{f>t\}}(x)dt.$$

Compare with:

$$f(x) = \int_0^\infty \mathbb{1}_{\{f>t\}}(x)dt.$$

Some properties of SDR

(1) $f^*(x)$ is radially symmetric decreasing.

(2)

$$\|f\|_p = \|f^*\|_p,\ \forall p > 0.$$

As a consequence, f, a density,

$$h_\alpha(f) = h_\alpha(f^*), \quad \forall \alpha \in (0, \infty).$$

For more properties and applications,
"Analysis", Lieb and Loss.
"A short course on rearrangement inequalities", Almut Burchard.

Measure the "peakedness" of distributions

Definition

For two probability densities f and g on $\mathbb{R}^d$, we say that f is majorized by g, written as $f \prec g$, if

$$\int_{\mathcal{B}(0,r)} f^*(x)\mathrm{d}x \leq \int_{\mathcal{B}(0,r)} g^*(x)\mathrm{d}x,$$

for all $r > 0$.

Measure the "peakedness" of distributions

Definition

For two probability densities f and g on $\mathbb{R}^d$, we say that f is majorized by g, written as $f \prec g$, if

$$\int_{\mathcal{B}(0,r)} f^*(x)\mathrm{d}x \leq \int_{\mathcal{B}(0,r)} g^*(x)\mathrm{d}x,$$

for all $r > 0$.

Lemma

Let f and g be two probability densities, with $f \prec g$, then

$$h_\alpha(f) \geq h_\alpha(g), \quad \forall \alpha \in (0, \infty).$$

Want:

$$h_\alpha(T(X) + W) \le h_\alpha(X + W)$$

$X \sim f_X$, $T(X) \sim f_{T(X)}$, it is straightforward to check that

$$f_X \prec f_{T(X)}.$$

Question:
Will convolution with a "nice" random variable W preserve majorization

$$f_{X+W} \prec f_{T(X)+W}?$$

A sample result: W radially symmetric log-concave, T any linear contraction, X log-concave

Theorem (AALMZ '22)

Let W be a radially symmetric, log-concave, $\mathbb{R}^d$-valued random variable. Then for any log-concave $\mathbb{R}^d$-valued random variable X that is independent of W, any linear contraction T, and $\alpha \in (0, \infty)$, we have

$$f_{X+W} \prec f_{T(X)+W}.$$

As a consequence,

$$h_\alpha(T(X) + W) \leq h_\alpha(X + W).$$

Corollary

Let $K \subset \mathbb{R}^d$ be a convex body. Then for any linear contaction T, one has

$$\mathrm{Vol}_d(T(K) + \mathcal{B}) \leq \mathrm{Vol}_d(K + \mathcal{B}).$$

Remark: Steiner's formula + Intrinsic volumes decrease under linear contractions[Grigoris Paouris, Peter Pivovarov, 2012].

$$V_i(T(K)) \leq V_i(K), \quad i = 1, \cdots, d.$$

A representation lemma

Lemma (Almut Burchard's notes)

Let $f : \mathbb{R}^d \to [0, \infty)$ *be an integrable non-negative function. Then,*

$$\int_{\mathcal{B}(0,r)} f^*(x)\mathrm{d}x = \sup_{\{C:\mathrm{Vol}(C)=\mathrm{Vol}(\mathcal{B}(0,r))\}} \int_C f(x)\mathrm{d}x.$$

Moreover, the supremum in the RHS is attained by the super-level set $\{f > t\}$ *having same volume as* $\mathcal{B}(0, r)$*.*

Sketch the proof of $f_{X+W} \prec f_{T(X)+W}$

$$X \sim f(x),\ W \sim g(x) \text{ and } T(X) \sim f_T(x).$$

Goal:

$$f \star g \prec (f_T) \star g$$

Note that

$$\{f \star g > t\},\ \text{bounded, convex}$$

Suffice to show that for any bounded convex set $K \subset \mathbb{R}^d$, $\exists K'$, having same volume as K, s.t.

$$\int_K (f \star g)(x)\mathrm{d}x \le \int_{K'} ((f_T) \star g)(x)\mathrm{d}x.$$

- 1 dimensional case. $T(x) = \lambda x$ for some $\lambda \in [0,1]$.
  suffice to prove

$$\int_K g(x-y)\mathrm{d}x \le \int_{K'} g(x-\lambda y)\mathrm{d}x.$$

Define

$$I(y) := \int_K g(x-y)\mathrm{d}x = \int_K g(y-x)\mathrm{d}x = \int_{K-y} g(x)\mathrm{d}x$$

$I(y)$ attains its maximum at y_0, the midpoint of K, and

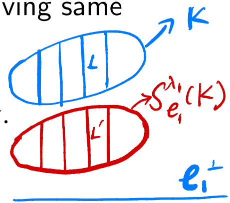
$$I((1-\lambda)y_0 + \lambda y) \geq I(y), \ \forall y \in \mathbb{R}.$$

This reads as

$$\int_{K-(1-\lambda)y_0} g(x-\lambda y)\mathrm{d}x \geq \int_K g(x-y)\mathrm{d}x.$$

Set $K' = K - (1-\lambda)y_0$.

- Higher dimensional case.
One may assume that T is diagonal, entries being $\lambda_1, \dots, \lambda_d \in [0,1]$.
Goal: For a bounded convex set $K \subset \mathbb{R}^d$, find K', having same volume as K, s.t.



$$\int_K g(x-y)\mathrm{d}x \le \int_{K'} g(x - T(y))\mathrm{d}x.$$

By Fubini's theorem,

$$\int_K g(x-y)\mathrm{d}x = \int_{K|_{e_1^\perp}} \left(\int_{L(x_2,\dots,x_n)} g(x_1 - y_1, x_2 - y_2, \dots, x_n - y_n)\mathrm{d}x_1 \right)$$

$$\le \int_{S_{e_1}^{\lambda_1}(K)} g(x_1 - \lambda_1 y_1, x_2 - y_2, \dots, x_n - y_n)\mathrm{d}x.$$

where,

$$S_{e_1}^{\lambda_1}(K) = \{y + \alpha e_1 : y \in K|_{e_1^\perp}, \alpha \in [g(y), f(y)] - (1-\lambda_1)\frac{f(y)+g(y)}{2}\}$$

Theorem (Anderson 1955)

Let K be an origin symmetric convex body in $\mathbb{R}^d$ and let g be a nonnegative, symmetric, unimodal,and integral function on $\mathbb{R}^d$. Then

$$\int_K g(x + cy)\mathrm{d}x \geq \int_K g(x + y)\mathrm{d}x,$$

for all $0 \leq c \leq 1$ and $y \in \mathbb{R}^d$.

Theorem (AALMZ '22)

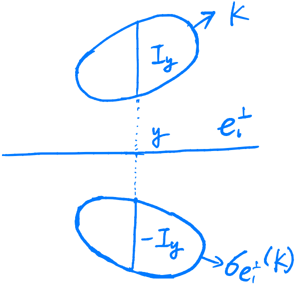
For any convex body K in $\mathbb{R}^d$, any diagonal matrix T with diagonal entries in $[0, 1]$, g radially symmetric log-concave, there exists a convex body $S(K)$ having same volume as K, such that

$$\int_{S(K)} g(x + T(y))\mathrm{d}x \geq \int_K g(x + y)\mathrm{d}x$$

for all $y \in \mathbb{R}^d$.

A different proof of $V_i(T(K)) \leq V_i(K)$

$$T = \begin{pmatrix} \lambda_1 & & 0 \\ & 1 & \\ 0 & & \ddots \ 1 \end{pmatrix}, \quad \lambda_1 \in [0,1]$$



$$K = \{y + \alpha e_1 : y \in K|_{e_1^\perp},\ g(y) \leq \alpha \leq f(y)\},$$

$$S_{e_1}^{\lambda_1}(K) = \{y+\alpha e_1 : y \in K|_{e_1^\perp}, \alpha \in \theta_{\lambda_1}[g(y), f(y)]+(1-\theta_{\lambda_1})[-f(y), -g(y)]\}$$

where $\theta_{\lambda_1} = \frac{1+\lambda_1}{2} \in [0,1]$.

$$\lambda_1 I_y \subseteq \theta_{\lambda_1} I_y + (1-\theta_{\lambda_1})(-I_y)$$

$$\Rightarrow \quad T(K) \subseteq S_{e_1}^{\lambda_1}(K)$$

(shadow systems of convex sets, Shephard '64)

$\lambda_1 \in [-1,1]$, $V_i(S_{e_1}^{\lambda_1}(K))$ is convex.

$$V_i(T(K)) \leq V_i(S_{e_1}^{\lambda_1}(K)) \leq V_i(K)$$

$\alpha = 2$, W radially symmetric log-concave, T any contraction, X arbitrary

Theorem (AALMZ '22)

Let X be an $\mathbb{R}^d$-valued random variable, and W be an independent $\mathbb{R}^d$-valued radially-symmetric log-concave random variable. Then, for any contraction $T : \mathbb{R}^d \to \mathbb{R}^d$, we have

$$h_2(T(X) + W) \leq h_2(X + W).$$

Metric distortion of Rényi entropies

$X = \{x_1, \cdots, x_n\}$, equipped with metric δ, for $\alpha \in [0,1) \cup (1,\infty)$, $p \in \Delta_n$, define:

$$D^\delta_\alpha(p) = \left(\sum_{i=1}^n \left(\sum_{j=1}^n e^{-\delta(x_i,x_j)} p_j \right)^{\alpha-1} p_i \right)^{1/(1-\alpha)}$$

$\lim_{t\to\infty} D^{t\delta}_\alpha(p) = e^{H_\alpha(p)}$

Diversities, introduced by Cobbold-Leinster ('12), developed further by Leinster-Meckes('15), Leinster-Roff ('19) , Aishwarya-Li-Madiman-Meckes ('22+).

Metric distortion of Rényi entropies

$X = \{x_1, \cdots, x_n\}$, equipped with metric δ, for $\alpha \in [0,1) \cup (1,\infty)$, $p \in \Delta_n$, define:

$$D_\alpha^\delta(p) = \left(\sum_{i=1}^{n} \left(\sum_{j=1}^{n} e^{-\delta(x_i,x_j)} p_j \right)^{\alpha-1} p_i \right)^{1/(1-\alpha)}$$

Diversities, introduced by Cobbold-Leinster ('12), developed further by Leinster-Meckes('15), Leinster-Roff ('19) , Aishwarya-Li-Madiman-Meckes ('22+).
(X, δ), $\mu \in \mathcal{P}(X)$,

$$D_\alpha^\delta(\mu) = \left(\int \left(\int e^{-\delta(x,y)} \mathrm{d}\mu(y) \right)^{\alpha-1} \mathrm{d}\mu(x) \right)^{1/(1-\alpha)}$$

Recovery of Renyi entropies in $\mathbb{R}^d$

Proposition (Aishwarya-Li-Madiman-Meckes '22+)

Assume that $\mathbb{R}^d$ is equipped with the Euclidean metric $\delta(x, y) = \|x - y\|_2$. Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ with density $f(x)$, then

$$e^{h_2(f)} = \lim_{t \to \infty} \frac{D_2^{t\delta}(\mu)}{c_d t^d},$$

where c_d is a constant depending on the dimension.

Sketch the proof of $h_2(T(X)+W) \le h_2(X+W)$

Suffice to show that

$$D_2^{t\delta}(X+W) \ge D_2^{t\delta}(T(X)+W).$$

X' i.i.d. X
W' i.i.d. W
Assume jointly independent

$$\left(D_2^{t\delta}(X+W)\right)^{-1} = \mathbb{E}e^{-t\|(X+W)-(X'+W')\|_2}.$$

$$= \mathbb{E}_{X,X'}\mathbb{E}_{W,W'}e^{-t\|(X-X')-(W'-W)\|_2}$$

Suffice to prove that for fixed x, x',

$$\mathbb{E}_{W,W'}e^{-t\|(x-x')-(W'-W)\|_2} \le \mathbb{E}_{W,W'}e^{-t\|(T(x)-T(x'))-(W'-W)\|_2}$$

Read as

$$\left(e^{-t\|\cdot\|_2} \star f_{(W'-W)}\right)(x-x') \le \left(e^{-t\|\cdot\|_2} \star f_{(W'-W)}\right)(T(x)-T(x'))$$

Recall that $\|T(x)-T(x')\|_2 \le \|x-x'\|_2$.

More open problems

X isotropic Gaussian, true

$T(x)=\lambda x$, $\lambda\in[0,1]$

T, linear, true

recover Costa's entropy power inequality (Costa '85)

(1) $N(X)=e^{\frac{2h_1(X)}{d}}$, $Z\sim\mathcal{N}(0,I_d)$, T be any contraction. Do we have

$$N(X+Z)\geq N(T(X)+Z)+(1-(\mathrm{Lip}(T))^2)N(X)?$$

(2) Let X be an arbitrary $\mathbb{R}^d$-valued log-concave random variable, T be any contraction and $Z\sim\mathcal{N}(0,I_d)$ be the standard Gaussian random variable. Can one show that

$$f_{X+\sqrt{t}Z}\prec f_{T(X)+\sqrt{t}Z},\quad \forall t>0?$$

Assume that the density of $T(X)$ exists, note that $f_X\prec f_{T(X)}$.

$$f_{e^{-t}X+\sqrt{1-e^{-2t}}Z}\prec f_{e^{-t}T(X)+\sqrt{1-e^{-2t}}Z}$$

Any Questions?