Hypercontractivity of spherical averages in Hamming space

Yury Polyanskiy

Department of EECS
MIT

yp@mit.edu

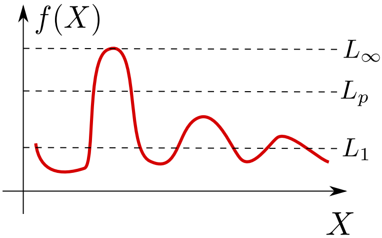
Allerton Conference, Oct 2, 2014

Hypercontractivity: definition

- L_p-norms of random variables:

$$\|f(x)\|_{L_p(\mu)} \triangleq \left(\int |f(x)|^p \, \mu(dx)\right)^{\frac{1}{p}}$$

$$\|U\|_p \triangleq (\mathbb{E}[|U|^p])^{\frac{1}{p}}$$



- Let P_{XY} – joint distribution.
Define $T_{Y|X}$ – stochastic (Markov) averaging operator

$$T_{Y|X} f(x) \triangleq \mathbb{E}[f(Y)|X = x]$$

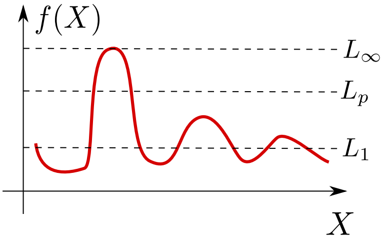
$\Longrightarrow$ $\|Tf\|_p \leq \|f\|_p$ by Jensen's.

Hypercontractivity: definition

- L_p-norms of random variables:

$$\|f(x)\|_{L_p(\mu)} \triangleq \left(\int |f(x)|^p \, \mu(dx)\right)^{\frac{1}{p}}$$

$$\|U\|_p \triangleq (\mathbb{E}[|U|^p])^{\frac{1}{p}}$$



- Let P_{XY} – joint distribution.
  Define $T_{Y|X}$ – stochastic (Markov) averaging operator

$$T_{Y|X} f(x) \triangleq \mathbb{E}[f(Y)|X = x]$$

$\implies \|Tf\|_p \leq \|f\|_p$ by Jensen's.

- T is hypercontractive if $\|T\|_{p \to q} = 1$ for $p < q$:

$$\|Tf\|_q \leq \|f\|_p \,, \qquad \|T\|_{p\to q} \triangleq \sup_f \frac{\|Tf\|_q}{\|f\|_p}$$

Tf gains integrability! (by smoothing peaks)

- T is HC for some $p \iff T$ is HC for all p.

Hypercontractivity: examples

- Nelson-Gross: (X, Y) – jointly Gaussian:

$$\|\mathbb{E}[f(Y)|X]\|_4 \le \|f(Y)\|_2 \quad \iff \quad \rho(X,Y) \le \frac{1}{\sqrt{3}}\,.$$

- Bonami-Gross: $(X, Y) \sim \frac{1}{2}\begin{bmatrix} 1-\delta & \delta \\ \delta & 1-\delta \end{bmatrix} \triangleq \mathrm{BSS}(\delta)$:

$$\|\mathbb{E}[f(Y)|X]\|_q \le \|f(Y)\|_p \quad \iff \quad \frac{p-1}{q-1} \ge (1-2\delta)^2$$

- Bonami-Beckner:

$$\|T\|_{p\to q} = 1 \quad \iff \quad \|T^{\otimes n}\|_{p\to q} = 1$$

- Ahslwede-Gács: for finite X, Y pretty much any $T_{Y|X}$ is HC

What is HC good for?

Canonical example:

- Estimate probability of a rectangle $\mathbb{P}[X \in A, Y \in B]$, where

$$\mathbb{P}[X \in A] \approx \mathbb{P}[Y \in B] = \theta \ll 1\,.$$

- Spectral gap (expanders, etc): $1 = \sigma_1(T) > \sigma_2(T) > \cdots$

$$\mathbb{P}[X \in A, Y \in B] \le \theta^2 + \sigma_2\theta(1-\theta) \approx \sigma_2 \cdot \theta$$

- Hypercontractivity:

$$\begin{aligned}\mathbb{P}[X \in A, Y \in B] &= (1_A, T1_B) \\ &\le \|1_A\|_{q'}\|T1_B\|_q \le \|1_A\|_{q'}\|1_B\|_p \\ &= \theta^{1-\frac{1}{q}+\frac{1}{p}} = \theta^{1+\epsilon} \ll \sigma_2\theta \quad (!)\end{aligned}$$

Note: Used $0/1$ nature of indicator functions.

Hypercontractivity and KL-divergence

$$
\begin{tikzcd}
P_X \arrow[dr] & & & P_Y \\
& X \arrow[r, "P_{Y|X}"] & Y \arrow[ur] \arrow[dr] & \\
Q_X \arrow[ur] & & & Q_Y
\end{tikzcd}
$$

- Apply to typical sets:

$$
\begin{aligned}
2^{-nD(Q_X\|P_X)} &\approx \mathbb{P}[X^n \in T^n_{[Q_X]}, Y^n \in T^n_{[Q_Y]}] \\
&\lesssim 2^{-n\frac{1}{q'}D(Q_X\|P_X) - n\frac{1}{p}D(Q_Y\|P_Y)}
\end{aligned}
$$

- So we get:

$$
\|T_{Y|X}\|_{p\to q} = 1 \quad \Longrightarrow \quad \boxed{D(Q_Y\|P_Y) \le \frac{p}{q} D(Q_X\|P_X) \quad \forall Q_X}
$$

Hypercontractivity and KL-divergence

$$\begin{tikzcd}
P_X \arrow[dr] & & & P_Y \\
& X \arrow[r, "P_{Y|X}"] & Y \arrow[ur] \arrow[dr] & \\
Q_X \arrow[ur] & & & Q_Y
\end{tikzcd}$$

- Apply to typical sets:

$$\begin{aligned}
2^{-nD(Q_X\|P_X)} &\approx \mathbb{P}[X^n \in T^n_{[Q_X]}, Y^n \in T^n_{[Q_Y]}] \\
&\lesssim 2^{-n\frac{1}{q'}D(Q_X\|P_X)-n\frac{1}{p}D(Q_Y\|P_Y)}
\end{aligned}$$

- So we get:

$$\|T_{Y|X}\|_{p\to q} = 1 \quad \underset{\not\Longleftarrow}{\Longrightarrow} \quad \boxed{D(Q_Y\|P_Y) \le \frac{p}{q} D(Q_X\|P_X) \quad \forall Q_X}$$

Hypercontractivity and KL-divergence

$$\begin{tikzcd}
P_X \arrow[dr] & & & P_Y \\
& X \arrow[r, "P_{Y|X}"] & Y \arrow[ur] \arrow[dr] & \\
Q_X \arrow[ur] & & & Q_Y
\end{tikzcd}$$

- For fixed P_{XY}, KL-contraction ratio:

$$\eta_{\mathrm{KL}}(P_{XY}) \triangleq \sup_{Q_X} \frac{D(Q_Y||P_Y)}{D(Q_X||P_X)} = \sup_{U \to X \to Y} \frac{I(U;Y)}{I(U;X)}$$

Hypercontractivity and KL-divergence

$$\begin{tikzcd}
P_X \arrow[dr] & & & P_Y \\
& X \arrow[r, "P_{Y|X}"] & Y \arrow[ur] \arrow[dr] & \\
Q_X \arrow[ur] & & & Q_Y
\end{tikzcd}$$

- For fixed P_{XY}, KL-contraction ratio:

$$\eta_{\mathrm{KL}}(P_{XY}) \triangleq \sup_{Q_X} \frac{D(Q_Y\|P_Y)}{D(Q_X\|P_X)} = \sup_{U \to X \to Y} \frac{I(U;Y)}{I(U;X)}$$

Theorem (Ahlswede-Gács)

$$\eta_{\mathrm{KL}}(P_{XY}) = \inf\left\{\frac{p}{q} : \|T_{Y|X}\|_{p\to q} = 1\right\}$$

- Roughly:

$$\forall Q_X : D(Q_Y\|P_Y) \le r D(Q_X\|P_X) \quad \Longrightarrow \boxed{\|T_{Y|X}\|_{p\to\frac{p}{r}} = 1, p \gg 1}$$

Hypercontractivity and KL-divergence

[Nair'2014]:

- Apply HC to typical sets of *arbitrary* Q_{XY} to get

$$\|T_{Y|X}\|_{p\to q} = 1 \quad \Longrightarrow$$

$$\boxed{\frac{1}{p}D(Q_Y\|P_Y) \le \frac{1}{q}D(Q_X\|P_X) + D(Q_{Y|X}\|P_{Y|X}|Q_X) \quad \forall Q_{XY}}$$

Hypercontractivity and KL-divergence

[Nair'2014]:

- Apply HC to typical sets of *arbitrary* Q_{XY} to get

$$\|T_{Y|X}\|_{p\to q} = 1 \quad \Longleftrightarrow$$

$$\boxed{\frac{1}{p}D(Q_Y\|P_Y) \le \frac{1}{q}D(Q_X\|P_X) + D(Q_{Y|X}\|P_{Y|X}|Q_X) \quad \forall Q_{XY}}$$

- Proof: Take $\|g_1\|_1 = 1$ and let $f = (g_1)^{\frac{1}{p}} \in L_p$
- STP: $\|Tf\|_q \le 1$.

Ahlswede-Gács:

$$Q_X(x) = P_X(x)\frac{Tf(x)^q}{\|Tf\|_q^q}$$

$$Q_{Y|X}(y|x) = P_{Y|X}(y|x)$$

Nair:

$Q_X(x) =$ ... same ...

$$Q_{Y|X}(y|x) = P_{Y|X}(y|x)\cdot\frac{f(y)}{Tf(x)}$$

$$\log\|Tf\|_q \le \frac{1}{p}D(Q_Y\|P_Y) - \frac{1}{q}D(Q_X\|P_X) - D(Q_{Y|X}\|P_{Y|X}|Q_X) \le 0$$

Hypercontractivity and KL-divergence

[Nair'2014]:

- Apply HC to typical sets of *arbitrary* Q_{XY} to get

$$\|T_{Y|X}\|_{p\to q} = 1 \quad \Longleftrightarrow$$

$$\boxed{\frac{1}{p}D(Q_Y\|P_Y) \le \frac{1}{q}D(Q_X\|P_X) + D(Q_{Y|X}\|P_{Y|X}|Q_X) \quad \forall Q_{XY}}$$

- Proof: Take $\|g_1\|_1 = 1$ and let $f = (g_1)^{\frac{1}{p}} \in L_p$
- STP: $\|Tf\|_q \le 1$.

Ahlswede-Gács:	Nair:
$Q_X(x) = P_X(x)\frac{Tf(x)^q}{\|Tf\|_q^q}$ $Q_{Y\|X}(y\|x) = P_{Y\|X}(y\|x)$	$Q_X(x) =$ ... same ... $Q_{Y\|X}(y\|x) = P_{Y\|X}(y\|x)\cdot\frac{f(y)}{Tf(x)}$

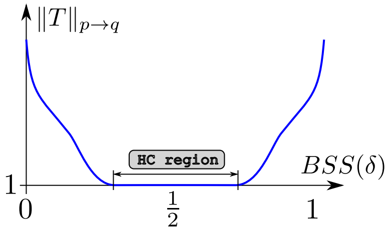
$$\log\|Tf\|_q \le \frac{1}{p}D(Q_Y\|P_Y) - \frac{1}{q}D(Q_X\|P_X) - D(Q_{Y|X}\|P_{Y|X}|Q_X) \le 0$$

- Note: For $p \gg 1$ both have $Q_{Y|X} \approx P_{Y|X}$.

$L_p \to L_q$ norm as measure of dependence

$\|T_{Y|X}\|_{p\to q}$ measures dependence:

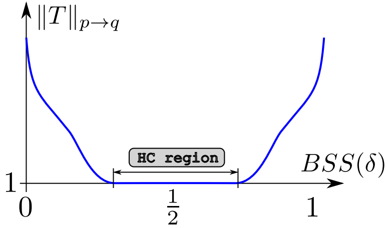
- Invariant to relabeling
- Satisfies data-processing
- Unusual: "sticky at 1".



$L_p \to L_q$ norm as measure of dependence

$\|T_{Y|X}\|_{p\to q}$ measures dependence:

- Invariant to relabeling
- Satisfies data-processing
- Unusual: "sticky at 1".



- Turns out the last one is a general phenomenon:

$\boxed{\text{All operators } T \text{ sufficiently noisy are } p \to q \text{ HC.}}$

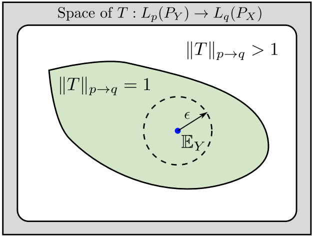
Semenov-Shneiberg'88

Theorem

For any $q \le 2 \le p$ *there is* $\epsilon = \epsilon(p,q)$ *s.t.*
simultaneously for all P_{XY}:

$$\|T_{Y|X} - \mathbb{E}_Y\|_{p\to q} \le \epsilon \quad \Rightarrow \quad \|T_{Y|X}\|_{p\to q} = 1$$

where $\mathbb{E}_Y f \triangleq 1_X \cdot \mathbb{E}[f(Y)]$.



Remarks:

- Also for general $q < p$ with extra condition: $\|T - \mathbb{E}\|_{2\to 2} \le \epsilon$
- Also for other operators of the type:

$$T1_Y = 1_X, \quad (1_X, T\cdot) = (1_X, \mathbb{E}\cdot)$$

- Orig. result is for $P_X = P_Y$, but proof easily generalizes.
- For finite (X,Y), suff. to check

$$\|P_{XY} - P_X P_Y\|_{TV} \le \epsilon' = \epsilon'(p, q, P_X, P_Y)$$

Proof for $q \le 2 \le p$:

- WLOG can take $f = 1 + Z$ with $Z \perp 1$. Decompose $T = \mathbb{E} + \Delta$. STP:

$$(*) \qquad \|1 + \Delta Z\|_q \le \|1 + Z\|_p$$

- Cool fact: $\exists$ universal τ, a, b s.t.

$$\|1 + Z\|_q \le 1 + a\|Z\|_q^2 \qquad \text{if } \|Z\|_q \le \tau$$
$$\|1 + Z\|_p \ge 1 + b\|Z\|_p^2 \qquad \text{if } \|Z\|_p \le \tau$$

- Whenever $\|\Delta\|_{p\to q}^2 \le \frac{b}{a}$ and $\|Z\|_p \le \tau$:

$$\|1 + \Delta Z)\|_q \le 1 + a\|\Delta Z\|_q^2 \le \|1 + Z\|_p$$

Proof for $q \le 2 \le p$:

- WLOG can take $f = 1 + Z$ with $Z \perp 1$. Decompose $T = \mathbb{E} + \Delta$. STP:
$$(*) \qquad \|1 + \Delta Z\|_q \le \|1 + Z\|_p$$
- Cool fact: $\exists$ universal τ, a, b s.t.
$$\|1 + Z\|_q \le 1 + a\|Z\|_q^2 \qquad \text{if } \|Z\|_q \le \tau$$
$$\|1 + Z\|_p \ge 1 + b\|Z\|_p^2 \qquad \text{if } \|Z\|_p \le \tau$$
- Whenever $\|\Delta\|_{p\to q}^2 \le \frac{b}{a}$ and $\|Z\|_p \le \tau$:
$$\|1 + \Delta Z)\|_q \le 1 + a\|\Delta Z\|_q^2 \le \|1 + Z\|_p$$
- Large Z: Another cool fact: $\exists$ universal $\beta_p(\cdot)$
$$\|1 + Z\|_p \ge 1 + \beta_p(\|Z\|_p)\,, \qquad \text{and } \frac{\beta_p(u)}{u}\text{–increasing}$$

  So (*) holds if $\|\Delta\|_{p\to q} \le \frac{\beta_p(\tau)}{\tau}$.

Spherical averages in Hamming space

- Hamming space $\mathbb{F}_2^n$ with uniform probability
- Bonami-Gross operator

$$N_\delta f(x) \triangleq \mathbb{E}[f(x+Z)], \qquad Z \sim \mathrm{Bern}(\delta)^n$$

- Spherical average:

$$T_\delta f(x) \triangleq \mathbb{E}[f(x+Z)], \qquad Z \sim \mathrm{Unif}(\mathbb{S}_{\delta n})$$

where $\mathbb{S}_j \triangleq \{y : \mathrm{wt}(y) = j\}$.

Theorem (P'13)

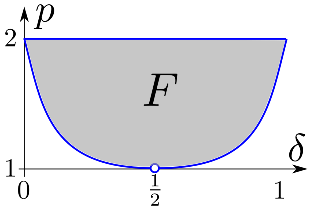
$\forall\, \delta \neq \frac{1}{2}$ and $p = 1 + (1-2\delta)^2$ there is a dim-independent $C_{p,\delta}$ s.t.

$$\|T_\delta f\|_2 \le C_{p,\delta} \|f\|_p$$

Remark: Bonami-Gross $\|N_\delta f\|_2 \le 1 \cdot \|f\|_p$

Hypercontractivity of T_δ: exact version

$$F = \left\{ (\delta, p) : \begin{array}{r} 1 + (1-2\delta)^2 \le p \le 2 \\ 0 \le \delta \le 1 \\ p \ne 1 \end{array} \right\}$$



Theorem (P'13)

For any compact subset $K \subset F$ there is a constant C_K s.t.

$$\|T_\delta\|_{p\to 2} \le C_K \qquad \forall(\delta, p) \in K\,,$$

where $T_\delta f = f * P_Z, P_Z = \mathrm{Unif}(\mathbb{S}_{\delta n})$.

For $p = 1, \delta = 1/2$*:*

$$\|T_{\frac{1}{2}}\|_{p\to 2} = \theta(n^{\frac{1}{4}})\,.$$

Note: for Bernoulli noise $\|N_\delta\|_{p\to 2} = 1$ everywhere on closure of F.

Proof: Prelim remarks

Why is constant not 1?

- Let $1_{even} = 1\{x : \mathrm{wt}(x)\text{-even}\}$
- Then:

$$T_\delta 1_{even} = 1_{even} \qquad \text{(or } 1_{odd}\text{)}$$

- So:

$$\|T_\delta\|_{p\to q} \ge 2^{\frac{1}{p}-\frac{1}{q}} > 1\,.$$

Why is constant not 1?

- Let $1_{even} = 1\{x : \mathrm{wt}(x)\text{-even}\}$
- Then:

$$T_\delta 1_{even} = 1_{even} \qquad \text{(or } 1_{odd}\text{)}$$

- So:

$$\|T_\delta\|_{p\to q} \ge 2^{\frac{1}{p}-\frac{1}{q}} > 1\,.$$

Beautiful methods that failed:

- Tensorization (T_δ not a tensor power)
- Semigroup method, aka $\frac{d}{d\delta}$ (T_δ not a semigroup)
- Comparison of Dirichlet forms (not useful for $\delta \gg \frac{1}{n}$)

Ugly method that worked:

- Direct computation of eigenvalues of T_δ

Proof: Orthogonal decomposition

- T_δ and N_δ are S_n-equivariant $\Rightarrow$

$$f(x) = \mathrm{poly}((-1)^{x_1}, \ldots, (-1)^{x_n}) = \sum_{a=0}^{n} f_a(x)\,,$$

  where $f_a(x) =$ degree a part of the poly.
- By odd/even trick can assume $f_a = 0$ for $a \geq n/2$.
- Have:

$$N_\delta f_a = (1-2\delta)^a \cdot f_a$$

$$T_\delta f_a = (\quad ?? \quad) \cdot f_a$$

Proof: Orthogonal decomposition

- T_δ and N_δ are S_n-equivariant $\Rightarrow$

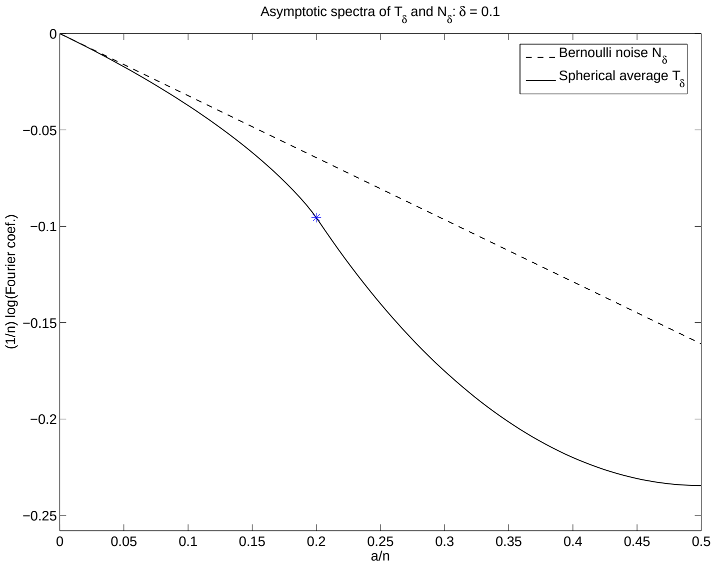
$$f(x) = \mathrm{poly}((-1)^{x_1}, \ldots, (-1)^{x_n}) = \sum_{a=0}^{n} f_a(x)\,,$$

  where $f_a(x) =$ degree a part of the poly.
- By odd/even trick can assume $f_a = 0$ for $a \geq n/2$.
- Have:

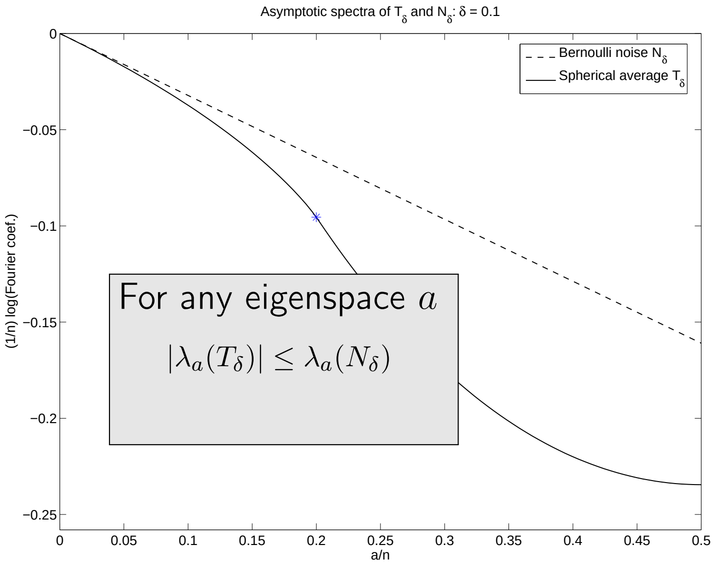
$$N_\delta f_a = (1-2\delta)^a \cdot f_a$$
$$T_\delta f_a = \kappa_{\delta n}(a) \cdot f_a$$

  and $\kappa_{\delta n}(a)$ – Krawtchouk polynomial of degree δn.

Proof (cont'd): Eigenvalues of N_δ vs T_δ



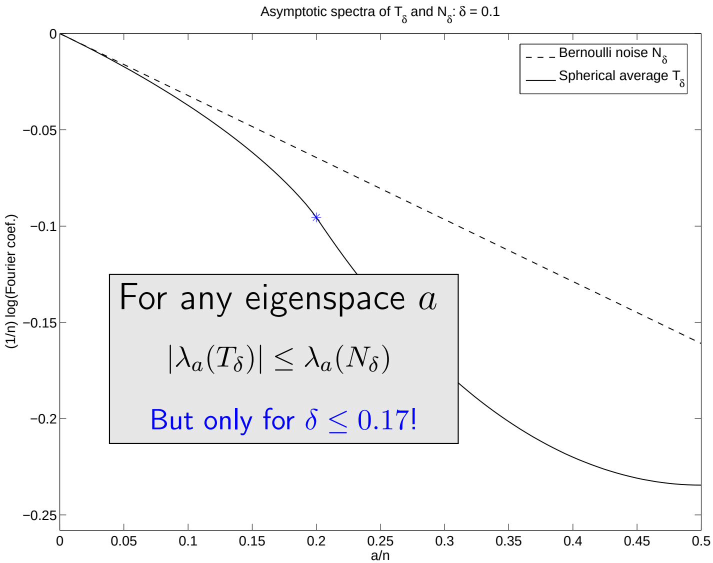
Proof (cont'd): Eigenvalues of N_δ vs T_δ



For any eigenspace a

$$|\lambda_a(T_\delta)| \le \lambda_a(N_\delta)$$

Proof (cont'd): Eigenvalues of N_δ vs T_δ

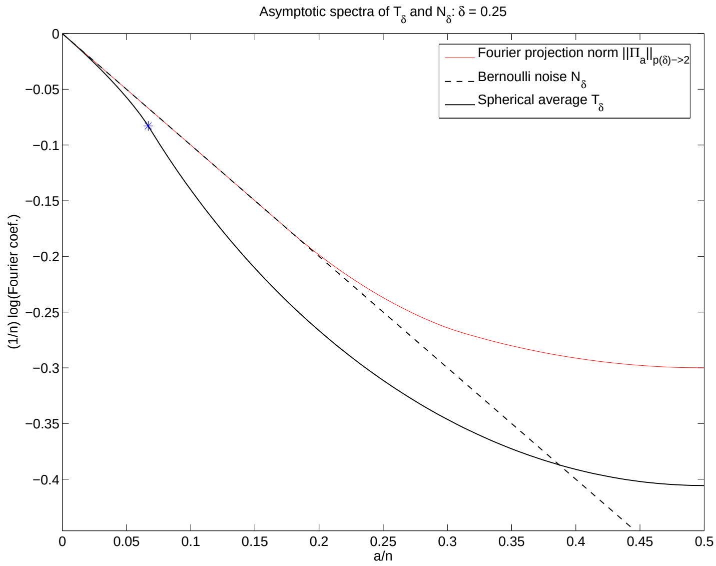


For any eigenspace a

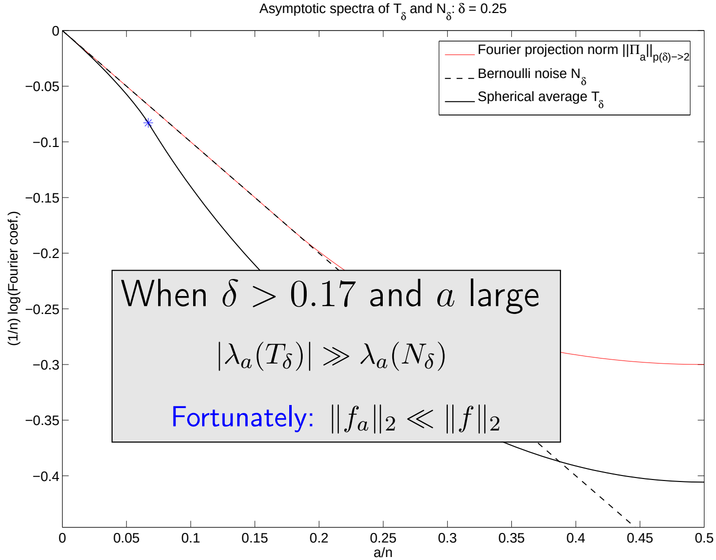
$$|\lambda_a(T_\delta)| \le \lambda_a(N_\delta)$$

But only for $\delta \le 0.17$!

Proof (cont'd): Failure for $\delta \gtrsim 0.17$



Proof (cont'd): Failure for $\delta \gtrsim 0.17$



When $\delta > 0.17$ and a large

$$|\lambda_a(T_\delta)| \gg \lambda_a(N_\delta)$$

Fortunately: $\|f_a\|_2 \ll \|f\|_2$

Proof (cont'd): Overall argument

- Know:

$$\|N_\delta f\|_2^2 = \sum_a (1-2\delta)^{2a}\|f_a\|_2^2 \le \|f\|_p^2 \qquad \text{by Bonami!}$$

- Need to show:

$$\|T_\delta f\|_2^2 = \sum_a \kappa_{\delta n}(a)^2 \|f_a\|_2^2 \le C\|f\|_p^2$$

- $a \le a_{crit}$:

$$\sum_{a \le a_{crit}} \kappa_{\delta n}(a)^2 \|f_a\|_2^2 \le C\sum_a (1-2\delta)^{2a}\|f_a\|_2^2$$

$$= C\|N_\delta f\|_2^2 \le C\|f\|_p^2$$

- $a > a_{crit}$: Need extra estimate:

$$\kappa_{\delta n}(a)\|f_a\|_2 \lesssim e^{-Cn}\|f\|_p$$

$\Longrightarrow$ $\sum_{a \ge \xi_{crit} n}$ contributes $\le C\|f\|_p^2$.

Proof (cont'd): Overall argument

- Know:

$$\|N_\delta f\|_2^2 = \sum_a (1-2\delta)^{2a} \|f_a\|_2^2 \le \|f\|_p^2 \qquad \text{by Bonami!}$$

- Need to show:

$$\|T_\delta f\|_2^2 = \sum_a \kappa_{\delta n}(a)^2 \|f_a\|_2^2 \le C \|f\|_p^2$$

- $a \le a_{crit}$: For $\delta < 0.17$: $a_{crit} = n/2$!

$$\sum_{a \le a_{crit}} \kappa_{\delta n}(a)^2 \|f_a\|_2^2 \le C \sum_a (1-2\delta)^{2a} \|f_a\|_2^2$$

$$= C\|N_\delta f\|_2^2 \le C\|f\|_p^2$$

- $a > a_{crit}$: Need extra estimate:

$$\kappa_{\delta n}(a) \|f_a\|_2 \lesssim e^{-Cn} \|f\|_p$$

$\Longrightarrow$ $\sum_{a \ge \xi_{crit} n}$ contributes $\le C\|f\|_p^2$.

Application to additive-combinatorics

Known fact:

Subspace $L \subset \mathbb{F}_2^n$ s.t. $|L \cap \mathbb{S}_{\delta n}| \sim |\mathbb{S}_{\delta n}|$ implies $\dim L \approx n$.

Application to additive-combinatorics

Known fact:

Subspace $L \subset \mathbb{F}_2^n$ s.t. $|L \cap \mathbb{S}_{\delta n}| \sim |\mathbb{S}_{\delta n}|$ implies $\dim L \approx n$.

Corollary

Let $A \subset \mathbb{F}_2^n$ be a subset s.t. sumset $A + A$ intersects some $\mathbb{S}_{\delta n}$ with average multiplicty $> \lambda|A|$. Then

$$|A| \geq \lambda^{\epsilon} \cdot 2^n .$$

Proof:

- Statement means

$$2^n(1_A * 1_A, 1_{\mathbb{S}_{\delta n}}) \geq \lambda \cdot |A| \cdot |\mathbb{S}_{\delta n}|$$

- Rearrange $(1_A * 1_A, 1_{\mathbb{S}}) = (T_\delta 1_A, 1A)$
- Apply Hölder and HC.

Thank you!

References

R. Ahlswede and P. Gacs, "Spreading of sets in product spaces and hypercontraction of the Markov operator," *Ann. Probab.*, pp. 925–939, 1976.

E. M. Semenov and I. Y. Shneiberg, "Hypercontractive operators and Khinchin's inequality," *Func. Analysis and Appl.*, vol. 22, no. 3, pp. 244–246, 1988.

Y. Polyanskiy, "Hypercontractivity of spherical averages in Hamming space," *arXiv:1309.3014*, Sep. 2013

Y. Polyanskiy, "Hypothesis testing via a comparator and hypercontractivity," *preprint*, Feb. 2013.

C. Nair, "Equivalent formulations of hypercontractivity using information measures", *Zurich Seminar on Comm.*, 2014

Dobrushin coefficient

- total variation

$$\|P - Q\|_{\mathrm{TV}} = \frac{1}{2}\int |\mathrm{d}P - \mathrm{d}Q|$$

- Dobrushin coefficient

$$\eta_{\mathrm{TV}} = \sup_{P_X \neq Q_X} \frac{\|P_Y - Q_Y\|_{\mathrm{TV}}}{\|P_X - Q_X\|_{\mathrm{TV}}} = \sup_{x,x'} \|P_{Y|X=x} - P_{Y|X=x'}\|_{\mathrm{TV}}.$$

- Original motivation: mixing of Markov chains

Consider a M.C. with invariant dist. P^*:

$$X_0 \to X_1 \to X_2 \to \cdots$$

Contraction:

$$\|P_{X_n} - P^*\|_{\mathrm{TV}} \le (\eta_{\mathrm{TV}})^n \qquad \text{(Dobrushin)}$$

$$\chi^2(P_{X_n}\|P^*) \le (\eta_{\chi^2})^n \cdot \chi^2(P_{X_0}\|P^*) \qquad \text{(spectral gap)}$$

$$D(P_{X_n}\|P^*) \le (\eta_{\mathrm{KL}})^n \cdot D(P_{X_0}\|P^*) \qquad \text{(log-Sobolev)}$$

Contraction coeffs and mixing of Markov chains

Consider a M.C. with invariant dist. P^*:

$$X_0 \to X_1 \to X_2 \to \cdots$$

Contraction:

$$\|P_{X_n} - P^*\|_{\mathrm{TV}} \le (\eta_{\mathrm{TV}})^n \qquad \text{(Dobrushin)}$$

$$\chi^2(P_{X_n} \| P^*) \le (\eta_{\chi^2})^n \cdot \chi^2(P_{X_0} \| P^*) \qquad \text{(spectral gap)}$$

$$D(P_{X_n} \| P^*) \le (\eta_{\mathrm{KL}})^n \cdot D(P_{X_0} \| P^*) \qquad \text{(log-Sobolev)}$$

TODO: add mixing estimate via HC

From TV to KL ... and further

Theorem (Cohen-Iwasa-Rautu-Ruskai-Seneta-Zbăganu'93)

$$\eta_{\mathrm{KL}} \le \eta_{\mathrm{TV}}$$

From TV to KL ... and further

Theorem (Cohen-Iwasa-Rautu-Ruskai-Seneta-Zbăganu'93)

$$\eta_{\mathrm{KL}} \le \eta_{\mathrm{TV}}$$

Fun application:

- Let T – kernel of a finite-state M.C. with invariant distribution P^*.
- Let $\mathbb{E}$ be indep. M.C.: $\mathbb{E}f \triangleq \int f dP^*$
- If $\|T(x,\cdot) - \mathbb{E}(x,\cdot)\|_{\mathrm{TV}} < 1/2$ then

$$\eta_{\mathrm{TV}} = \max_{x,x'} \|T(x,\cdot) - T(x',\cdot)\|_{\mathrm{TV}} < 1$$

- Then $\eta_{\mathrm{KL}} < 1$. By Ahslwede-Gács $\exists q \gg 1$

$$\|Tf\|_q \le \|f\|_p, \qquad p = \eta_{\mathrm{KL}} \cdot q < q \qquad \text{(hypercont)}$$

- $\Rightarrow$ There is a ball around $\mathbb{E}$ s.t. for all T: $\|T\|_{L_p \to L_q} = 1$.

Theorem (Cohen-Iwasa-Rautu-Ruskai-Seneta-Zbăganu'93)

$$\eta_{\mathrm{KL}} \le \eta_{\mathrm{TV}}$$

Fun application:

- Let T – kernel of a finite-state M.C. with invariant distribution P^*.
- Let $\mathbb{E}$ be indep. M.C.: $\mathbb{E}f \triangleq \int f dP^*$
- If $\|T(x,\cdot) - \mathbb{E}(x,\cdot)\|_{\mathrm{TV}} < 1/2$ then

$$\eta_{\mathrm{TV}} = \max_{x,x'} \|T(x,\cdot) - T(x',\cdot)\|_{\mathrm{TV}} < 1$$

- Then $\eta_{\mathrm{KL}} < 1$. By Ahslwede-Gács $\exists q \gg 1$

$$\|Tf\|_q \le \|f\|_p, \qquad p = \eta_{\mathrm{KL}} \cdot q < q \qquad \text{(hypercont)}$$

- $\Rightarrow$ There is a ball around $\mathbb{E}$ s.t. for all T: $\|T\|_{L_p \to L_q} = 1$.
- $\Rightarrow$ Every (mixing) M.C. eventually (hyper-)contracts $L_p \to L_q$.

Theorem (Cohen-Iwasa-Rautu-Ruskai-Seneta-Zbăganu'93)

$$\eta_{\mathrm{KL}} \le \eta_{\mathrm{TV}}$$

Fun application:

- Let T – kernel of a finite-state M.C. with invariant distribution P^*.
- Let $\mathbb{E}$ be indep. M.C.: $\mathbb{E}f \triangleq \int f dP^*$
- If $\|T(x,\cdot) - \mathbb{E}(x,\cdot)\|_{\mathrm{TV}} < 1/2$ then

$$\eta_{\mathrm{TV}} = \max_{x,x'} \|T(x,\cdot) - T(x',\cdot)\|_{\mathrm{TV}} < 1$$

- Then $\eta_{\mathrm{KL}} < 1$. By Ahslwede-Gács $\exists q \gg 1$

$$\|Tf\|_q \le \|f\|_p, \qquad p = \eta_{\mathrm{KL}} \cdot q < q \qquad \text{(hypercont)}$$

- $\Rightarrow$ There is a ball around $\mathbb{E}$ s.t. for all T: $\|T\|_{L_p \to L_q} = 1$.
- $\Rightarrow$ Every (mixing) M.C. eventually (hyper-)contracts $L_p \to L_q$.
- Euclidean space: planar face of the ball of Fourier multipliers: [Segal'70], [Fefferman-Shapiro'72], [Semenov-Shneiberg'88]