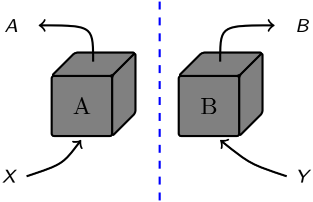
Optimizing conditional entropies for quantum correlations

Omar Fawzi

ECM, Computational aspects of commutative and noncommutative positive polynomials

Based on joint works with Peter Brown and Hamza Fawzi
arXiv:2007.12575 and arXiv:2106.13692

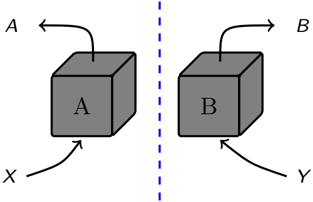
Bell-nonlocality



- Defines a conditional distribution $p(ab|xy)$
- Noncommutative polynomial optimization (NPA hierarchy): decide if $p(ab|xy) \in \mathrm{Q}$
  $\mathrm{Q} = \{p(ab|xy) : \exists$ quantum strategy achieving $p\}$

Device-independent quantum cryptography

Bell-nonlocality

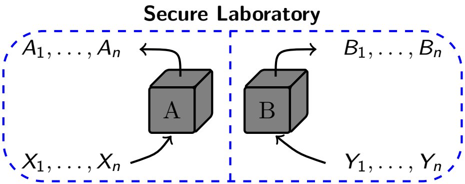


- Defines a conditional distribution $p(ab|xy)$
- Noncommutative polynomial optimization (NPA hierarchy): decide if
  $p(ab|xy) \in \mathrm{Q}$
  $\mathrm{Q} = \{p(ab|xy) : \exists$ quantum strategy achieving $p\}$
- Nonlocal correlations $\implies$ randomness in the outcomes
- Foundation for device-independent protocols (key distribution, randomness expansion,...)
- **This talk:** For the analysis of protocols, want to compute more complicated properties related to quantum strategies

Device-independent randomness expansion

A randomness expansion protocol

1. Choose $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ at random
   w.p. γ, set $X_i Y_i \sim \mu(x, y)$ and w.p. $1 - \gamma$ set $X_i = x^*$ and $Y_i = y^*$
2. Device interaction

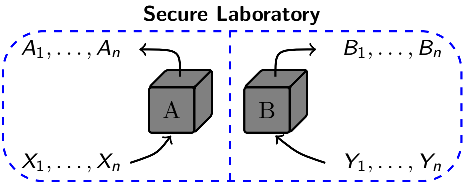


3. From the outputs $A_1, \ldots, A_n, B_1, \ldots, B_n$, estimate $p(ab|xy)$ (for one round)
4. If $p(ab|xy)$ is sufficiently non-local, extract the randomness by applying f
   $f(A_1, \ldots, A_n, B_1, \ldots, B_n) \in \{0, 1\}^\ell$

Question: How large can we take ℓ?

A randomness expansion protocol

1. Choose $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ at random
   w.p. γ, set $X_i Y_i \sim \mu(x,y)$ and w.p. $1-\gamma$ set $X_i = x^*$ and $Y_i = y^*$
2. Device interaction

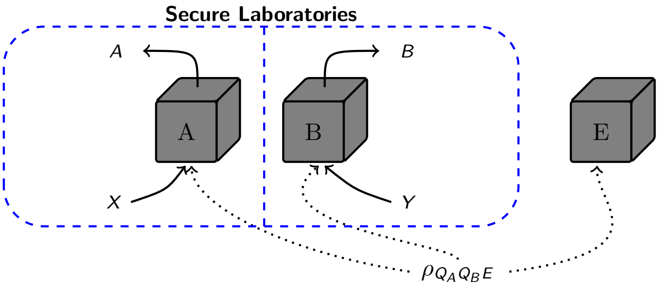


3. From the outputs $A_1, \ldots, A_n, B_1, \ldots, B_n$, estimate $p(ab|xy)$ (for one round)
4. If $p(ab|xy)$ is sufficiently non-local, extract the randomness by applying f
   $f(A_1, \ldots, A_n, B_1, \ldots, B_n) \in \{0,1\}^\ell$

Question: How large can we take ℓ?

One can show

$$\ell = n \times \text{randomness generated by device compatible with } p - O(\sqrt{n})$$



A **strategy** (i.e., implementation of the boxes) is a tuple

$$(Q_A \otimes Q_B \otimes Q_E, \rho_{Q_A Q_B E}, \{\{M_{a|x}\}_a\}_x, \{\{N_{b|y}\}_b\}_y)$$

With each strategy we can associate a **post-measurement state**

$$\rho_{ABXYE} = \sum_{abxy} \mu(xy)|abxy\rangle\langle abxy| \otimes \mathrm{tr}_{Q_A Q_B}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right] .$$

It is **compatible** with $p(ab|xy)$ if

$$p(ab|xy) = \mathrm{tr}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right] \qquad \forall x, y, a, b \ .$$

Randomness generated per round

A **strategy** (i.e., implementation of the boxes) is a tuple

$$(Q_A \otimes Q_B \otimes Q_E, \rho_{Q_A Q_B E}, \{\{M_{a|x}\}_a\}_x, \{\{N_{b|y}\}_b\}_y)$$

With each strategy we can associate a **post-measurement state**

$$\rho_{ABXYE} = \sum_{abxy} \mu(xy)|abxy\rangle\langle abxy| \otimes \mathrm{tr}_{Q_A Q_B}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right].$$

$$\boxed{\begin{aligned} \text{rand. gen. per round} = \inf_{\text{strategies}} \quad & H(AB|X = x^*, Y = y^* E)_{\rho_{ABXYE}} \\ \text{s.t.} \quad & \mathrm{tr}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right] = p(ab|xy) \quad \forall xyab \end{aligned}}$$

For a quantum state ρ_{AE},

$$H(A|E)_\rho = -\mathrm{tr}\left[\rho_{AE} \log \rho_{AE}\right] + \mathrm{tr}\left[\rho_E \log \rho_E\right] \ ,$$

where $\rho_E = \mathrm{tr}_A[\rho_{AE}]$.

Objective: Good lower bounds
Difficulty: Objective function is not linear in $\rho_{Q_A Q_B E}$

Trial 1: Approximate log with powers

Rényi conditional entropy with $\alpha \in (1,2]$

$$H(A|E)_\rho \geq H_\alpha(A|E)_\rho = \frac{1}{1-\alpha} \log \operatorname{tr}\left[\rho_{AE}^{\alpha} \rho_{E}^{1-\alpha}\right] \ ,$$

Objective becomes:

$$\boxed{\begin{array}{ll} \sup\limits_{\text{strategies}} & \sum\limits_{ab} \operatorname{tr}\left[\left(\operatorname{tr}_{Q_A Q_B}\left[(M_{a|x^*} \otimes N_{b|y^*} \otimes I_E)\rho_{Q_A Q_B E}\right]\right)^{\alpha} \rho_E^{1-\alpha}\right] \\ \text{s.t.} & \operatorname{tr}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right] = p(ab|xy) \end{array}}$$

Replacing the logarithm with powers: Rényi entropy

Trial 1: Approximate log with powers

Rényi conditional entropy with $\alpha \in (1,2]$

$$H(A|E)_\rho \geq H_\alpha(A|E)_\rho = \frac{1}{1-\alpha} \log \mathrm{tr}\left[\rho_{AE}^{\alpha} \rho_E^{1-\alpha}\right] ,$$

Objective becomes:

$$\boxed{\begin{array}{ll} \sup\limits_{\text{strategies}} & \sum\limits_{ab} \mathrm{tr}\left[\left(\mathrm{tr}_{Q_A Q_B}\left[(M_{a|x^*} \otimes N_{b|y^*} \otimes I_E)\rho_{Q_A Q_B E}\right]\right)^{\alpha} \rho_E^{1-\alpha}\right] \\ \text{s.t.} & \mathrm{tr}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right] = p(ab|xy) \end{array}}$$

Difficulties: Handle partial trace? Rational powers?

Dimension-free variational expressions

Trial 2: Use specific properties of $H(A|E)_\rho$

Fact: $\rho \mapsto H(A|E)_\rho$ is *concave* $\implies \exists \mathcal{F}$ s.t. $H(A|E)_\rho = \inf_{(Z,z)\in\mathcal{F}} \mathrm{tr}[\rho Z] + z$

Recall the problem

$$\begin{aligned} \inf_{\text{strategies}} \quad & H(AB|X = x^*, Y = y^* E)_{\rho_{ABXYE}} \\ \text{s.t.} \quad & \mathrm{tr}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right] = p(ab|xy) \end{aligned}$$

$$\begin{aligned} H(AB|X = x^*, Y = y^* E)_\rho &= \inf_{(Z,z)\in\mathcal{F}} \mathrm{tr}\left[\rho_{ABXYE} Z\right] + z \\ &= \inf_{(Z,z)\in\mathcal{F}} \sum_{a,b} \mathrm{tr}\left[\mathrm{tr}_{Q_A Q_B}[M_{a|x^*} \otimes N_{b|y^*} \rho_{Q_A Q_B E}] \langle ab|Z|ab\rangle\right] + z \\ &= \inf_{(Z,z)\in\mathcal{F}} \sum_{a,b} \mathrm{tr}\left[\rho_{Q_A Q_B E} M_{a|x^*} \otimes N_{b|y^*} I_{Q_A Q_B} \otimes \langle ab|Z|ab\rangle\right] + z \end{aligned}$$

If $\mathcal{F}$ is described by dimension-free polynomial constraints
$\implies$ can use NC poly optimization machinery

Approaches to obtain dimension-free variational expressions

We proposed two methods to give dimension-free variational lower bounds on $H(A|E)_\rho$

1. Based on SDP representations of the **matrix geometric mean**
   [Nat Commun 12, 575 (2021)]
   Rényi entropy: $\mathrm{tr}[\rho_{AE}^{\alpha}\rho_E^{1-\alpha}] \leq \mathrm{tr}[\rho_{AE}\#_{1-\alpha}(I_A \otimes \rho_E)]$

$$X\#Y := X^{1/2}(X^{-1/2}YX^{-1/2})^{1/2}X^{1/2} = \max\left\{W : \begin{pmatrix} X & W \\ W & Y \end{pmatrix} \geq 0\right\}$$

   Use this idea to **define** new Rényi entropies (iterated mean) $\leq H(A|E)$

Approaches to obtain dimension-free variational expressions

We proposed two methods to give dimension-free variational lower bounds on $H(A|E)_\rho$

1. Based on SDP representations of the **matrix geometric mean**
   [Nat Commun 12, 575 (2021)]
   Rényi entropy: $\mathrm{tr}[\rho_{AE}^{\alpha}\rho_{E}^{1-\alpha}] \leq \mathrm{tr}[\rho_{AE}\#_{1-\alpha}(I_A \otimes \rho_E)]$

$$X\#Y := X^{1/2}(X^{-1/2}YX^{-1/2})^{1/2}X^{1/2} = \max\left\{W : \begin{pmatrix} X & W \\ W & Y \end{pmatrix} \geq 0\right\}$$

   Use this idea to **define** new Rényi entropies (iterated mean) $\leq H(A|E)$

2. Based on approximating log via rational functions
   [Soon on arXiv]
   Rest of talk: focus on this approach

Dimension-free variational expressions via rational functions

Note that $H(A|E)_\rho = -D(\rho_{AE} \| I_A \otimes \rho_E)$
where $D(\rho\|\sigma) = \mathrm{tr}[\rho \log \rho] - \mathrm{tr}[\rho \log \sigma]$

From now: work with the divergence D (called quantum relative entropy or Umegaki divergence)

Property: For any ρ, σ, there exists a measure $\nu_{\rho,\sigma}$ on $\mathbb{R}_+^2$ such that

$$D(\rho\|\sigma) = \int_{\mathbb{R}_+^2} y \log(y/x) d\nu_{\rho,\sigma}(x,y)$$

For $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ and $\sigma = \sum_k q_k |\phi_k\rangle\langle\phi_k|$, then $\nu_{\rho,\sigma} = \sum_{j,k} |\langle\psi_j|\phi_k\rangle|^2 \delta_{q_k,p_j}$

Dimension-free variational expressions via rational functions

Note that $H(A|E)_\rho = -D(\rho_{AE} \| I_A \otimes \rho_E)$
where $D(\rho\|\sigma) = \mathrm{tr}[\rho \log \rho] - \mathrm{tr}[\rho \log \sigma]$

From now: work with the divergence D (called quantum relative entropy or Umegaki divergence)

Property: For any ρ, σ, there exists a measure $\nu_{\rho,\sigma}$ on $\mathbb{R}_+^2$ such that

$$D(\rho\|\sigma) = \int_{\mathbb{R}_+^2} y \log(y/x) d\nu_{\rho,\sigma}(x,y)$$

For $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ and $\sigma = \sum_k q_k |\phi_k\rangle\langle\phi_k|$, then $\nu_{\rho,\sigma} = \sum_{j,k} |\langle\psi_j|\phi_k\rangle|^2 \delta_{q_k,p_j}$

Approximate log by a sum of rational functions via Gauss-Radau quadrature:

$$\ln(z) = \int_0^1 \frac{z-1}{t(z-1)+1} dt \geq \sum_{i=1}^{m} w_i \frac{z-1}{t_i(z-1)+1}$$

for some well-chosen nodes $t_i \in (0,1]$ and weights $w_i > 0$
Approximation gets arbitrary good as $m \to \infty$

Dimension-free variational expressions via rational functions

Note that $H(A|E)_\rho = -D(\rho_{AE} \| I_A \otimes \rho_E)$
where $D(\rho\|\sigma) = \mathrm{tr}[\rho \log \rho] - \mathrm{tr}[\rho \log \sigma]$

From now: work with the divergence D (called quantum relative entropy or Umegaki divergence)

Property: For any ρ, σ, there exists a measure $\nu_{\rho,\sigma}$ on $\mathbb{R}_+^2$ such that

$$D(\rho\|\sigma) = \int_{\mathbb{R}_+^2} y \log(y/x) d\nu_{\rho,\sigma}(x,y)$$

For $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ and $\sigma = \sum_k q_k |\phi_k\rangle\langle\phi_k|$, then $\nu_{\rho,\sigma} = \sum_{j,k} |\langle\psi_j|\phi_k\rangle|^2 \delta_{q_k,p_j}$

Approximate log by a sum of rational functions via Gauss-Radau quadrature:

$$\ln(z) = \int_0^1 \frac{z-1}{t(z-1)+1} dt \geq \sum_{i=1}^m w_i \frac{z-1}{t_i(z-1)+1}$$

for some well-chosen nodes $t_i \in (0,1]$ and weights $w_i > 0$
Approximation gets arbitrary good as $m \to \infty$

$$(\ln 2) D(\rho\|\sigma) = -\int_{\mathbb{R}_+^2} y \ln(x/y) d\nu_{\rho,\sigma}(x,y) \leq -\sum_{i=1}^m w_i \int_{\mathbb{R}_+^2} \frac{y(x-y)}{t_i(x-y)+y} d\nu_{\rho,\sigma}(x,y)$$

Dimension-free variational expressions for rational functions

Want a variational expression for the "rational function divergence"

$$D_t(\rho\|\sigma) := \int_{\mathbb{R}_+^2} \frac{y(x-y)}{t(x-y)+y} d\nu_{\rho,\sigma}(x,y)$$

$\frac{y(x-y)}{t(x-y)+y} = \frac{1}{t}\frac{1}{(t(x-y))^{-1}+y^{-1}} = \frac{1}{t}M_{-1}(t(x-y),y)$ with M_{-1} is the harmonic mean

There exists a vector v and operators A, B on some Hilbert space such that

$$D_t(\rho\|\sigma) = \langle v, M_{-1}(t(A-B),B)v\rangle$$

$A =$ left multiplication by σ and $B =$ right multiplication by ρ

Var. expression for harmonic mean: $\langle v, M_{-1}(X,Y)v\rangle = \inf_{z+z'=v}\langle z, Xz\rangle + \langle z', Yz'\rangle$

$$\begin{aligned} D_t(\rho\|\sigma) &= \inf_z \langle z, t(A-B)z\rangle + \langle v-z, B(v-z)\rangle \\ &= \inf_z t\ \mathrm{tr}[z^*\sigma z] - t\ \mathrm{tr}[z^* z\rho] + \mathrm{tr}[(v-z)^*(v-z)\rho] \\ &= \inf_z t\ \mathrm{tr}[zz^*\sigma] + (1-t)\mathrm{tr}[z^* z\rho] + \mathrm{tr}[\rho] - \mathrm{tr}[(z+z^*)\rho] \end{aligned}$$

using the fact that $v = I$

Back to the quantum relative entropy

Putting things together

$$(\ln 2)D(\rho\|\sigma) \leq -\sum_{i=1}^{m} w_i D_{t_i}(\rho\|\sigma)$$

$$\leq \boxed{-\inf_{z_1,\ldots,z_m} \sum_{i=1}^{m} w_i \left(t_i \ \mathrm{tr}[z_i z_i^* \sigma] + (1-t_i)\mathrm{tr}[z_i^* z_i \rho] + \mathrm{tr}[\rho] - \mathrm{tr}[(z_i + z_i^*)\rho]\right)}$$

Exactly the form we wanted when $m \to \infty$, we get equality

Back to the quantum relative entropy

Putting things together

$$(\ln 2)D(\rho\|\sigma) \leq -\sum_{i=1}^{m} w_i D_{t_i}(\rho\|\sigma)$$

$$\leq \boxed{-\inf_{z_1,\ldots,z_m} \sum_{i=1}^{m} w_i \left(t_i \ \mathrm{tr}[z_i z_i^* \sigma] + (1-t_i)\mathrm{tr}[z_i^* z_i \rho] + \mathrm{tr}[\rho] - \mathrm{tr}[(z_i + z_i^*)\rho]\right)}$$

Exactly the form we wanted when $m \to \infty$, we get equality

Back to motivating problem

$$\inf_{\text{strategies}} \quad H(AB|X = x^*, Y = y^* E)_{\rho_{ABXYE}}$$

$$\text{s.t.} \quad \mathrm{tr}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right] = p(ab|xy)$$

Apply formula for $\rho \leftarrow \rho_{ABE}$ and $\sigma \leftarrow I_{AB} \otimes \rho_E$ (all conditioned on $X = x^*, Y = y^*$)

$$H(AB|X = x^*, Y = y^* E)_\rho$$

$$\geq \inf_{Z_1,\ldots,Z_m} \sum_{i=1}^{m} w_i \left(1 + \mathrm{tr}\left[\rho_{ABE}(Z_i + Z_i^* + (1-t_i)Z_i^* Z_i)\right] + t \ \mathrm{tr}[I_{AB} \otimes \rho_E Z_i Z_i^*]\right)$$

$$= \inf_{Z_{i,ab}} \sum_{i=1}^{m} w_i \left(1 + \sum_{ab} \mathrm{tr}\left[\rho_{Q_A Q_B E} M_{a|x^*} N_{b|y^*} (Z_{i,ab} + Z_{i,ab}^* + (1-t_i)Z_{i,ab}^* Z_{i,ab})\right] + t_i \mathrm{tr}[\rho_E Z_{i,ab} Z_{i,ab}^*]\right)$$

The family of NC poly optimization

Back to motivating problem

$$\begin{aligned} \inf_{\text{strategies}} \quad & H(AB|X = x^*, Y = y^* E)_{\rho_{ABXYE}} \\ \text{s.t.} \quad & \operatorname{tr}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right] = p(ab|xy) \end{aligned}$$

Parameter $m \geq 1$

$$\begin{aligned} \inf_{Z_{i,ab}, M_{a|x}, N_{b|y}, |\psi\rangle} \quad & \sum_{i=1}^{m} w_i \left(1 + \sum_{ab} \langle\psi| M_{a|x^*} N_{b|y^*} (Z_{i,ab} + Z_{i,ab}^* + (1 - t_i) Z_{i,ab}^* Z_{i,ab}) |\psi\rangle + t_i \langle\psi| Z_{i,ab} Z_{i,ab}^* |\psi\rangle\right) \\ \text{s.t.} \quad & \langle\psi| M_{a|x} N_{b|y} |\psi\rangle = p(ab|xy) \\ & \sum_a M_{a|x} = \sum_b N_{b|y} = I \quad M_{a|x}, N_{b|y} \geq 0 \\ & [M_{a|x}, N_{b|y}] = 0 \\ & Z_{i,ab}, Z_{i,ab}^* \text{ commute with all } M_{a'|x}, N_{b'|y} \end{aligned}$$

Back to motivating problem

$$\begin{aligned} \inf_{\text{strategies}} \quad & H(AB|X=x^*, Y=y^*E)_{\rho_{ABXYE}} \\ \text{s.t.} \quad & \mathrm{tr}\left[(M_{a|x} \otimes N_{b|y} \otimes I_E)\rho_{Q_A Q_B E}\right] = p(ab|xy) \end{aligned}$$

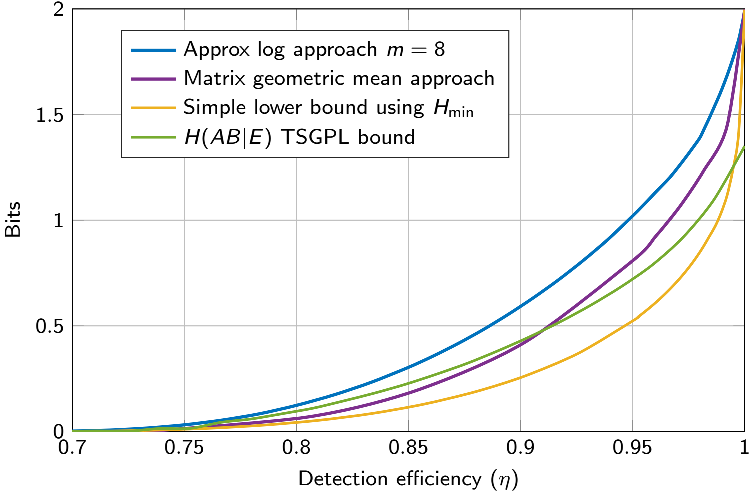
Parameter $m \geq 1$

$$\begin{aligned} \inf_{Z_{i,ab}, M_{a|x}, N_{b|y}, |\psi\rangle} \quad & \sum_{i=1}^{m} w_i \left(1 + \sum_{ab} \langle\psi| M_{a|x^*} N_{b|y^*} (Z_{i,ab} + Z_{i,ab}^* + (1-t_i) Z_{i,ab}^* Z_{i,ab}) |\psi\rangle + t_i \langle\psi| Z_{i,ab} Z_{i,ab}^* |\psi\rangle \right) \\ \text{s.t.} \quad & \langle\psi| M_{a|x} N_{b|y} |\psi\rangle = p(ab|xy) \\ & \sum_a M_{a|x} = \sum_b N_{b|y} = I \quad M_{a|x}, N_{b|y} \geq 0 \\ & [M_{a|x}, N_{b|y}] = 0 \\ & Z_{i,ab}, Z_{i,ab}^* \text{ commute with all } M_{a'|x}, N_{b'|y} \end{aligned}$$

Can give an a priori bound $\|Z_{i,ab}\| \leq \alpha_m$ to get convergence of NPA

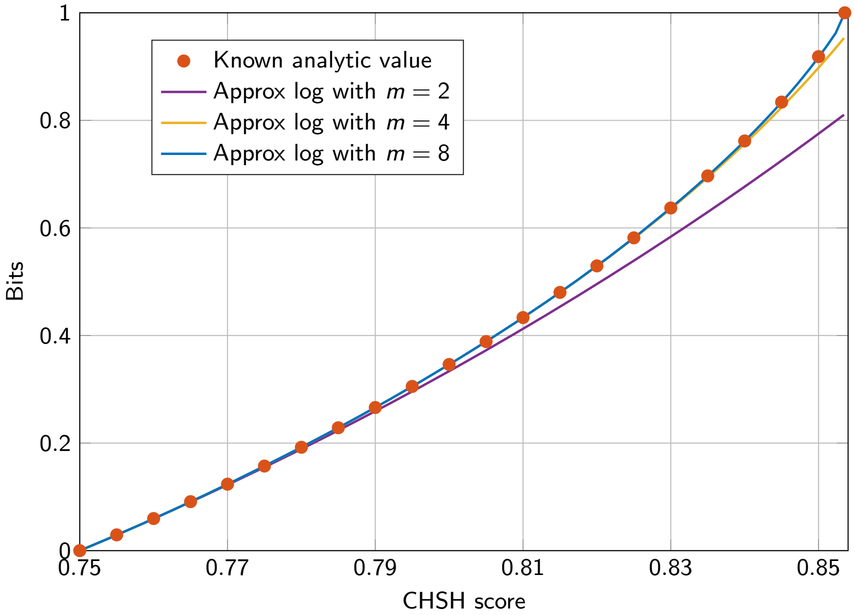
But the bound α_m grows with m

Application: randomness expansion



Numerically: Tighter bounds and computationally faster that other methods

A sample plot showing tightness



Open problems

- **Concrete problem about convergence:**
  NC polynomial p with variables X_j and Z_i with $[X_j, Z_i] = 0$
  Assume X_j all bounded and Z_i unrestricted
  Can one show convergence of SDP hierarchies?
- **More general methods:** We used concavity of entropy, can one construct hierarchies for more general settings? e.g., **maximizing** entropy?