Quantum relative entropy optimization

Hamza Fawzi

Department of Applied Mathematics and Theoretical Physics
University of Cambridge

based on joint works with James Saunderson (Monash Univ.),
and with Peter Brown (Telecom Paris) and Omar Fawzi (INRIA Lyon)

CWI, Amsterdam, September 2022

Entropy

Classical information theory If $p \geq 0$,

- Entropy $H(p) = -\sum_{i=1}^{n} p_i \log p_i$ (**Concave**).
- Kullback-Leibler divergence (or relative entropy)

$$D(p\|q) = \sum_{i=1}^{n} p_i \log(p_i/q_i)$$

Convex jointly in (p, q).

Matrix logarithm function

- X symmetric matrix with positive eigenvalues (positive definite)

$$X = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T \quad \rightarrow \quad \log(X) = U \begin{pmatrix} \log(\lambda_1) & & \\ & \ddots & \\ & & \log(\lambda_n) \end{pmatrix} U^T$$

where U orthogonal.

Matrix logarithm function

- X symmetric matrix with positive eigenvalues (positive definite)

$$X = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T \quad \to \quad \log(X) = U \begin{pmatrix} \log(\lambda_1) & & \\ & \ddots & \\ & & \log(\lambda_n) \end{pmatrix} U^T$$

where U orthogonal.

- Matrix (von Neumann) entropy of X:

$$H(X) = -\operatorname{Tr}[X \log X] = -\sum_{i=1}^{n} \lambda_i(X) \log(\lambda_i(X))$$

Concave in X. (Spectral function)

Matrix logarithm function

- X symmetric matrix with positive eigenvalues (positive definite)

$$X = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T \quad \rightarrow \quad \log(X) = U \begin{pmatrix} \log(\lambda_1) & & \\ & \ddots & \\ & & \log(\lambda_n) \end{pmatrix} U^T$$

  where U orthogonal.

- Matrix (von Neumann) entropy of X:

$$H(X) = -\operatorname{Tr}[X \log X] = -\sum_{i=1}^{n} \lambda_i(X) \log(\lambda_i(X))$$

  Concave in X. (Spectral function)

- (Umegaki) quantum relative entropy:

$$\boxed{D(X \| Y) = \operatorname{Tr}[X(\log X - \log Y)]}$$

  Convex in (X, Y) [Lieb-Ruskai, 1973]. Cornerstone result in quantum information. ($D(X\|Y)$ *not* spectral function!)

Applications

- Many problems in quantum information involve the quantum relative entropy function (quantum cryptography, computing quantum channel capacities, quantum many-body systems, ...)

  - Convex:

$$\min_{X,Y} \; D(X\|Y) \quad \text{s.t. } \mathcal{A}(X,Y)=b, \;\; X,Y \succeq 0.$$

Applications

- Many problems in quantum information involve the quantum relative entropy function (quantum cryptography, computing quantum channel capacities, quantum many-body systems, ...)

  - Convex:

$$\min_{X,Y} \; D(X\|Y) \quad \text{s.t. } \mathcal{A}(X,Y) = b, \;\; X, Y \succeq 0.$$

  - Nonconvex:

$$\max_{X,Y} \; D(X\|Y) \text{ s.t. } \mathcal{A}(X,Y) = b, \;\; X, Y \succeq 0.$$

Outline

This talk: algorithmic tools to solve quantum relative entropy optimization problems

1. Review of convexity of $D(X\|Y)$
2. Convex problems: a self-concordant barrier for the quantum relative entropy cone with optimal parameter
3. Nonconvex problems: a method to derive semidefinite relaxations

Outline

This talk: algorithmic tools to solve quantum relative entropy optimization problems

1. **Review of convexity of $D(X\|Y)$**

2. Convex problems: a self-concordant barrier for the quantum relative entropy cone with optimal parameter

3. Nonconvex problems: a method to derive semidefinite relaxations

Operator relative entropy

- Belavkin-Staszewski operator relative entropy

$$D_{op}(X\|Y) = X^{1/2}\log(X^{1/2}Y^{-1}X^{1/2})X^{1/2}$$

- Jointly operator convex in (X, Y), i.e.,

$$\mathrm{epi}(D_{op}) = \{(X, Y, T) : D_{op}(X\|Y) \preceq T\}$$

is a convex set.

Integral representation of log

$$\log(Y) = \int_0^1 \left(1 - \frac{1}{1+s(Y-1)}\right) \frac{ds}{s}$$

- Consequence:

$$D_{op}(X\|Y) = \int_0^1 \psi_s(X,Y) ds/s$$

where

$$\psi_s(X,Y) = X\left[(1-s)X + sY\right]^{-1} X - X$$

is jointly operator convex, since

$$\psi_s(X,Y) \preceq T \iff \underbrace{\begin{bmatrix} (1-s)X + sY & X \\ X & X+T \end{bmatrix} \succeq 0}_{\text{linear matrix inequality}}.$$

- $\implies$ $D_{op}(X\|Y)$ is convex as an (infinite) sum of convex functions.

Umegaki relative entropy

- Key observation [Pusz-Woronowicz, Ando, Petz, Effros, ...]:

$$D(X\|Y) = \phi(D_{op}(X \otimes I\|I \otimes Y))$$

where $\phi(Z) = \langle \mathrm{vec}(I_n), Z\mathrm{vec}(I_n)\rangle$ is a positive linear map.
$\implies D(X\|Y)$ is convex

Umegaki relative entropy

- Key observation [Pusz-Woronowicz, Ando, Petz, Effros, ...]:

$$D(X\|Y) = \phi(D_{op}(X \otimes I \| I \otimes Y))$$

  where $\phi(Z) = \langle \mathrm{vec}(I_n), Z\mathrm{vec}(I_n)\rangle$ is a positive linear map.
  $\implies D(X\|Y)$ is convex

- Interpretation:
  - $X \otimes I \in \mathbb{R}^{n^2 \times n^2}$, seen as a linear operator $\mathcal{R}_X : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, corresponds to right multiplication by X, i.e., $\mathcal{R}_X(a) = aX$.

Umegaki relative entropy

- Key observation [Pusz-Woronowicz, Ando, Petz, Effros, ...]:

$$D(X\|Y) = \phi(D_{op}(X \otimes I\|I \otimes Y))$$

  where $\phi(Z) = \langle \text{vec}(I_n), Z\text{vec}(I_n)\rangle$ is a positive linear map.
  $\implies D(X\|Y)$ is convex

- Interpretation:
  - $X \otimes I \in \mathbb{R}^{n^2 \times n^2}$, seen as a linear operator $\mathcal{R}_X : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, corresponds to right multiplication by X, i.e., $\mathcal{R}_X(a) = aX$.
  - Similarly $I \otimes Y \in \mathbb{R}^{n^2 \times n^2}$ corresponds to left multiplication by Y, $\mathcal{L}_Y : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, $\mathcal{L}_Y(a) = Ya$.

Umegaki relative entropy

- Key observation [Pusz-Woronowicz, Ando, Petz, Effros, ...]:

$$D(X\|Y) = \phi(D_{op}(X \otimes I \| I \otimes Y))$$

where $\phi(Z) = \langle \mathrm{vec}(I_n), Z\mathrm{vec}(I_n) \rangle$ is a positive linear map.
$\implies D(X\|Y)$ is convex

- Interpretation:
  - $X \otimes I \in \mathbb{R}^{n^2 \times n^2}$, seen as a linear operator $\mathcal{R}_X : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, corresponds to right multiplication by X, i.e., $\mathcal{R}_X(a) = aX$.
  - Similarly $I \otimes Y \in \mathbb{R}^{n^2 \times n^2}$ corresponds to left multiplication by Y, $\mathcal{L}_Y : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, $\mathcal{L}_Y(a) = Ya$.
  - $\mathcal{R}_X, \mathcal{L}_Y$ are self-adjoint (wrt Hilbert-Schmidt inner product), positive, and commute

$$D(X\|Y) = \langle I, D_{op}(\mathcal{R}_X \| \mathcal{L}_Y)(I) \rangle_{HS}.$$

1. Review of convexity of $D(X\|Y)$

2. **Convex problems: a self-concordant barrier for the quantum relative entropy cone with optimal parameter**

3. Nonconvex problems: a method to derive semidefinite relaxations

Semidefinite approximations

- In previous work [Fawzi, Saunderson, Parrilo] we derived semidefinite approximations of the quantum relative entropy function:

$$\underline{r}_m(X\|Y) \leq D(X\|Y) \leq \bar{r}_m(X\|Y)$$

  where $\underline{r}_m$ and $\bar{r}_m$ have a semidefinite representation, and converge to $D(X\|Y)$.

- Drawback: size of the semidefinite representation is $\approx mn^2$ $(X, Y \in \mathbf{S}^n)$

Semidefinite approximations

- In previous work [Fawzi, Saunderson, Parrilo] we derived semidefinite approximations of the quantum relative entropy function:

$$\underline{r}_m(X\|Y) \leq D(X\|Y) \leq \bar{r}_m(X\|Y)$$

  where $\underline{r}_m$ and $\bar{r}_m$ have a semidefinite representation, and converge to $D(X\|Y)$.

- Drawback: size of the semidefinite representation is $\approx mn^2$ $(X, Y \in \mathbf{S}^n)$

- Goal: optimize over the quantum relative entropy cone without paying the quadratic dependence on n

- $\rightarrow$ Interior-point methods for the quantum relative entropy cone

Interior-point methods

Conic programming over $\mathcal{K}$ (a convex cone)

$$\min\ \langle c, x \rangle \quad \text{s.t.} \quad Ax = b,\ x \in \mathcal{K}.$$

- A *barrier function* for $\mathcal{K}$ is a strictly convex function $F : \mathrm{int}(\mathcal{K}) \to \mathbb{R}$ such that $F(x) \to +\infty$ as $x \to \partial\mathcal{K}$.

- Interior-point methods follow *central path*

$$x^*(t) = \underset{t}{\operatorname{argmin}}\{t\,\langle c, x \rangle + F(x) \ \text{ s.t. } \ Ax = b\}$$

  as $t \to \infty$, via Newton's method.

- Nesterov-Nemirovski: if F is self-concordant (s.c.) with parameter ν, then interior-point method solves conic program in $\approx \sqrt{\nu}\log(1/\epsilon)$ Newton steps (where ϵ is the desired objective function accuracy).

Self-concordant functions

- Self-concordance is a condition that governs the variation of $\nabla^2 F(x)$ wrt the local norm induced by F

$$\|h\|_x = \sqrt{\langle h, \nabla^2 F(x) h \rangle}$$

Self-concordant functions

- Self-concordance is a condition that governs the variation of $\nabla^2 F(x)$ wrt the local norm induced by F

$$\|h\|_x = \sqrt{\langle h, \nabla^2 F(x) h \rangle}$$

- Formally, F is self-concordant if

$$\boxed{|D^3 F(x)[h]| \leq 2\|h\|_x^{3/2} \qquad \forall x \in \mathrm{dom}(F), h \in \mathbb{R}^n}$$

where $D^3 F(x)[h] = \frac{d^3}{dt^3} F(x + th)|_{t=0}$.

Self-concordant functions

- Self-concordance is a condition that governs the variation of $\nabla^2 F(x)$ wrt the local norm induced by F

$$\|h\|_x = \sqrt{\langle h, \nabla^2 F(x) h \rangle}$$

- Formally, F is self-concordant if

$$\boxed{|D^3 F(x)[h]| \leq 2\|h\|_x^{3/2} \qquad \forall x \in \mathrm{dom}(F), h \in \mathbb{R}^n}$$

where $D^3 F(x)[h] = \frac{d^3}{dt^3} F(x + th)|_{t=0}$.

- If F is self-concordant, then λF for $\lambda \geq 1$ is, and so is $x \mapsto F(Ax + b)$.

Self-concordant barriers

- Typical examples of barriers:
  - $\mathcal{K} = \mathbb{R}^n_+$: $F(x) = -\sum_{i=1}^n \log x_i$
  - $\mathcal{K} = \mathbf{S}^n_+$: $F(X) = -\log\det X$

- Goal: construct a s.c. barrier for

$$\mathcal{K}_{qre} = \{(X, Y, t) : D(X\|Y) \leq t\} = \text{epigraph of } D.$$

Main theorem

Theorem (Fawzi, Saunderson)

The function

$$F(X, Y, t) = -\log(t - D(X\|Y)) - \log\det X - \log\det Y$$

is a (logarithmically homogeneous) self-concordant barrier for

$$\mathcal{K}_{qre} = \{(X, Y, t) \in (\mathbf{S}_+^n)^2 \times \mathbb{R} : D(X\|Y) \leq t\}$$

of optimal parameter $2n+1$.

- Logarithmically homogeneous with parameter ν:

$$F(\lambda x) = F(x) - \nu \log \lambda$$

- "Natural" logarithmic barrier, first proposed by [Karimi, Tunçel]

Compatibility condition of Nesterov

- Assume $\psi : \mathrm{dom}(\psi) \to \mathbf{S}^m$ is *matrix convex* so that

$$\mathrm{epi}(\psi) = \{(Z, T) : \psi(Z) \preceq T\}$$

  is a convex set.

- A natural candidate for a self-concordant barrier of $\mathrm{epi}(\psi)$ is

$$(Z, T) \mapsto -\log\det(T - \psi(Z)) + G(Z)$$

  where G is a s.c. barrier of $\mathrm{dom}(\psi)$

Compatibility condition of Nesterov

- Assume $\psi : \mathrm{dom}(\psi) \to \mathbf{S}^m$ is *matrix convex* so that

$$\mathrm{epi}(\psi) = \{(Z, T) : \psi(Z) \preceq T\}$$

  is a convex set.

- A natural candidate for a self-concordant barrier of $\mathrm{epi}(\psi)$ is

$$(Z, T) \mapsto -\log\det(T - \psi(Z)) + G(Z)$$

  where G is a s.c. barrier of $\mathrm{dom}(\psi)$

- Nesterov: This is true provided the following *compatibility condition* is true:

$$\boxed{D^3\psi(z)[h] \;\; \preceq \;\; 3\,(D^2G(z)[h])^{1/2}\; D^2\psi(z)[h]} \quad \forall z \in \mathrm{dom}(\psi), h$$

Compatibility condition of Nesterov

- Assume $\psi : \mathrm{dom}(\psi) \to \mathbf{S}^m$ is *matrix convex* so that

$$\mathrm{epi}(\psi) = \{(Z, T) : \psi(Z) \preceq T\}$$

  is a convex set.
- A natural candidate for a self-concordant barrier of $\mathrm{epi}(\psi)$ is

$$(Z, T) \mapsto -\log\det(T - \psi(Z)) + G(Z)$$

  where G is a s.c. barrier of $\mathrm{dom}(\psi)$
- Nesterov: This is true provided the following *compatibility condition* is true:

$$\boxed{D^3\psi(z)[h] \;\; \preceq \;\; 3\,(D^2G(z)[h])^{1/2}\; D^2\psi(z)[h]} \quad \forall z \in \mathrm{dom}(\psi), h$$

- Key fact about condition: it is linear in ψ!
  - If ψ_1, ψ_2 compatible with G, then so is $\psi = \psi_1 + \psi_2$
  - Note: it is not in general easy to get a s.c. barrier for $\mathrm{epi}(\psi_1 + \psi_2)$ from s.c. barriers of $\mathrm{epi}(\psi_1)$ and $\mathrm{epi}(\psi_2)$.

Proof

- Recall
$$D(X\|Y) = \int_0^1 \phi(\psi_s(X \otimes I, I \otimes Y))ds/s$$
where ψ_s is a "nice" rational function.

- One can check "by hand" that the integrand
$$(X, Y) \mapsto \psi_s(X \otimes I, I \otimes Y)$$
satisfies the compatibility condition wrt $F(X, Y) = -\log\det X - \log\det Y$

- Thus $D(X\|Y)$ satisfies the compatibility condition too, and this proves that
$$(X, Y, t) \mapsto -\log(t - D(X\|Y)) - \log\det X - \log\det Y$$
is a s.c. barrier for $\mathcal{K}_{qre}$.

General result

- Löwner: any operator convex function $f : (0, \infty) \to \mathbb{R}$ has an integral representation similar to the one for logarithm.

Theorem (Fawzi, Saunderson)

If $f : (0, \infty) \to \mathbb{R}$ is operator convex and $P_f(X, Y) = X^{1/2} f(X^{-1/2} Y X^{-1/2}) X^{1/2}$ is its matrix perspective, then

$$F(X, Y, T) = -\log\det(T - P_f(X, Y)) - \log\det X - \log\det Y$$

is a s.c. barrier for $\mathcal{K}_f = \mathrm{epi}(P_f)$.

- Similar theorem can be proved for the epigraph of

$$(X, Y) \mapsto \phi(P_f(X \otimes I, I \otimes Y))$$

  where ϕ is any positive linear map.
- Allows us to get s.c. barriers for the quantum Rényi entropies

$$\mathrm{tr}[X^\alpha Y^{1-\alpha}], \ \ X^{1/2}(X^{-1/2} Y X^{-1/2})^\alpha X^{1/2}$$

Discussion

- Similar proof technique was used by Faybusovich and Zhou to obtain self-concordant barriers for $X \mapsto \mathrm{tr}(C \log(X))$ (i.e., one of the arguments of $D(X\|Y)$ is fixed).

- Basic path-following method of Nesterov & Nemirovski is too slow in practice. Primal-dual predictor-corrector methods are much faster. Can we use recent techniques used e.g., in MOSEK's exponential cone solver?

1. Review of convexity of $D(X\|Y)$

2. Convex problems: a self-concordant barrier for the quantum relative entropy cone with optimal parameter

3. **Nonconvex problems: a method to derive semidefinite relaxations**

Nonconvex relative entropy optimization

- We consider nonconvex problems, where the goal is to maximize the quantum relative entropy, e.g.,

$$\max_{X \succeq 0, \operatorname{tr} X = 1} D(\mathcal{A}(X) \| \mathcal{B}(X))$$

  where $\mathcal{A}, \mathcal{B}$ are two positive maps (i.e., $\mathcal{A}(X) \succeq 0, \ \forall X \succeq 0$).

- Nonpolynomial problem! Cannot use sum-of-squares/NPA, etc.

- **Goal:** derive semidefinite relaxations for this problem

Main idea

- Since $D(X\|Y)$ is convex and 1-homogeneous, it must have a variational formulation of the form

$$D(X\|Y) = \max_{(M,N)\in\mathcal{C}} \langle X, M\rangle + \langle Y, N\rangle$$

  for some complicated set $\mathcal{C}$

Main idea

- Since $D(X\|Y)$ is convex and 1-homogeneous, it must have a variational formulation of the form

$$D(X\|Y) = \max_{(M,N)\in\mathcal{C}} \langle X, M\rangle + \langle Y, N\rangle$$

for some complicated set $\mathcal{C}$

- We will see such a formula holds and takes the more precise form

$$\boxed{D(X\|Y) = \max_{a} \langle X, P(a)\rangle + \langle Y, Q(a)\rangle}$$

where P, Q are quadratic polynomials

Main idea

- Since $D(X\|Y)$ is convex and 1-homogeneous, it must have a variational formulation of the form

$$D(X\|Y) = \max_{(M,N)\in\mathcal{C}} \langle X, M\rangle + \langle Y, N\rangle$$

  for some complicated set $\mathcal{C}$

- We will see such a formula holds and takes the more precise form

$$\boxed{D(X\|Y) = \max_a \langle X, P(a)\rangle + \langle Y, Q(a)\rangle}$$

  where P, Q are quadratic polynomials

- Our original problem can thus be written as

$$\max_X D(\mathcal{A}(X)\|\mathcal{B}(X)) = \boxed{\max_{X,a} \mathrm{tr}[XS(a)]}$$

  where $S(a) = \mathcal{A}^*(P(a)) + \mathcal{B}^*(Q(a))$ is a polynomial.

Integral representation

Recall integral representation of $D(X\|Y)$:

$$D(X\|Y) = \int_0^1 \underbrace{\phi(\psi_s(X \otimes I, I \otimes Y))}_{D_s(X\|Y)} ds/s$$

Key fact: each $D_s(X\|Y)$ has a simple variational formulation.

Integral representation

Recall integral representation of $D(X\|Y)$:

$$D(X\|Y) = \int_0^1 \underbrace{\phi(\psi_s(X \otimes I, I \otimes Y))}_{D_s(X\|Y)} ds/s$$

Key fact: each $D_s(X\|Y)$ has a simple variational formulation.

Proposition

$$D_s(X\|Y) = \sup_{a \in \mathbb{R}^{n\times n}} \langle X, P_s(a)\rangle + \langle Y, Q_s(a)\rangle$$

where

$$\begin{cases} P_s(a) = -(I + a + a^T + (1-s)a^T a) \\ Q_s(a) = -saa^T \end{cases}$$

Proof of variational formula for D_s

$$\psi_s(X, Y) = X((1-s)X + sY)^{-1}X - X$$

Key ingredient:

- Schur complement corresponds to partial minimization of a quadratic form: for any block operator $\left[\begin{smallmatrix} A & B \\ B^T & C \end{smallmatrix}\right] \succ 0$ and any x,

$$\inf_y \left\langle \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle = \left\langle x, (A - BC^{-1}B^T)x \right\rangle.$$

$\implies$ Allows us to write $D_s(X\|Y) = \langle I, \psi_s(\mathcal{R}_X, \mathcal{L}_Y)(I)\rangle_{HS}$ as a supremum.

The variational formula for $D(X\|Y)$

$$\begin{aligned} D(X\|Y) &= \int_0^1 D_s(X\|Y) ds/s \\ &= \int_0^1 \sup_{a \in \mathbb{R}^{n\times n}} \langle X, P_s(a)\rangle + \langle Y, Q_s(a)\rangle \, ds/s \\ &= \sup_{\boldsymbol{a}=\{a(s)\}} \langle X, P(\boldsymbol{a})\rangle + \langle Y, Q(\boldsymbol{a})\rangle \end{aligned}$$

where

$$P(\boldsymbol{a}) = \int_0^1 P_s(a(s)) ds/s, \qquad Q(\boldsymbol{a}) = \int_0^1 Q_s(a(s)) ds/s.$$

This formula appeared in the works of [Kosaki, 1986] and [Donald, 1986]

Problem: infinite number of variables!

Discretizing the integral

$$\log(y) = \int_0^1 \underbrace{\frac{1}{s}\left(1 - \frac{1}{1+s(y-1)}\right)}_{f(s;y)} ds$$

Discretizing the integral

$$\log(y) = \int_0^1 \underbrace{\frac{1}{s}\left(1 - \frac{1}{1+s(y-1)}\right)}_{f(s;y)} ds \overset{?}{\approx} \sum_{i=1}^{m} w_i f(s_i; y) = r_m(y)$$

Discretizing the integral

$$\log(y) = \int_0^1 \underbrace{\frac{1}{s}\left(1 - \frac{1}{1+s(y-1)}\right)}_{f(s;y)} ds \overset{?}{\approx} \sum_{i=1}^{m} w_i f(s_i;y) = r_m(y)$$

- Gaussian quadrature: choose $s_1, \ldots, s_m$ and weights $w_1, \ldots, w_m > 0$ such that
$$\int_0^1 p(s)ds = \sum_{i=1}^{m} w_i p(s_i) \;\; \forall \deg(p) \leq 2m - 1.$$
Resulting rational approximation $r_m(y)$ is the diagonal (m, m) Padé approximation of the log.

Discretizing the integral

$$\log(y) = \int_0^1 \underbrace{\frac{1}{s}\left(1 - \frac{1}{1+s(y-1)}\right)}_{f(s;y)} ds \overset{?}{\approx} \sum_{i=1}^{m} w_i f(s_i; y) = r_m(y)$$

- Gaussian quadrature: choose $s_1, \ldots, s_m$ and weights $w_1, \ldots, w_m > 0$ such that
$$\int_0^1 p(s)ds = \sum_{i=1}^{m} w_i p(s_i) \;\; \forall \deg(p) \leq 2m-1.$$
Resulting rational approximation $r_m(y)$ is the diagonal (m, m) Padé approximation of the log.
- Gauss-Radau quadrature with one node fixed at $s = 1$:
$$\int_0^1 p(s)ds = \sum_{i=1}^{m-1} w_i p(s_i) + w_m p(1) \qquad \forall \deg(p) \leq 2m-2.$$
Resulting rational function $r_m(y)$ is a lower bound on $\log(y)$.

Putting things together

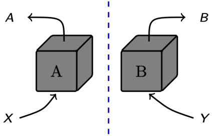
$$D(X\|Y) \le D^{(m)}(X\|Y) := \sup_{\boldsymbol{a}=(a_1,\ldots,a_m)} \left\langle X, P^{(m)}(\boldsymbol{a}) \right\rangle + \left\langle Y, Q^{(m)}(\boldsymbol{a}) \right\rangle$$

where

$$\begin{cases} P^{(m)}(\boldsymbol{a}) = \sum_{i=1}^{m} w_i P_{s_i}(a_i) \\ Q^{(m)}(\boldsymbol{a}) = \sum_{i=1}^{m} w_i Q_{s_i}(a_i). \end{cases}$$

Right-hand side converges to $D(X\|Y)$ as $m \to \infty$.

Quantum random number generators

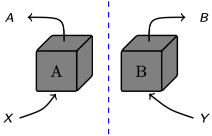


- Goal: generate shared randomness between two parties
- Amount of randomness generated, secure against adversary:

$$\begin{array}{ll} \inf\limits_{M,N,\rho} & H(A|E)_\rho \\ \text{s.t.} & \text{observed statistics (nc polynomial constraints).} \end{array}$$

where $H(A|E)$ is the conditional entropy.

Quantum random number generators



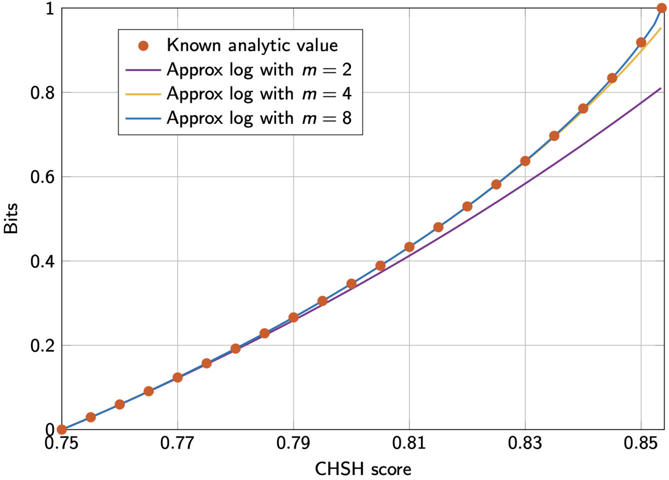
- Goal: generate shared randomness between two parties
- Amount of randomness generated, secure against adversary:

$$\begin{array}{ll} \inf\limits_{M,N,\rho} & H(A|E)_\rho \\ \text{s.t.} & \text{observed statistics (nc polynomial constraints).} \end{array}$$

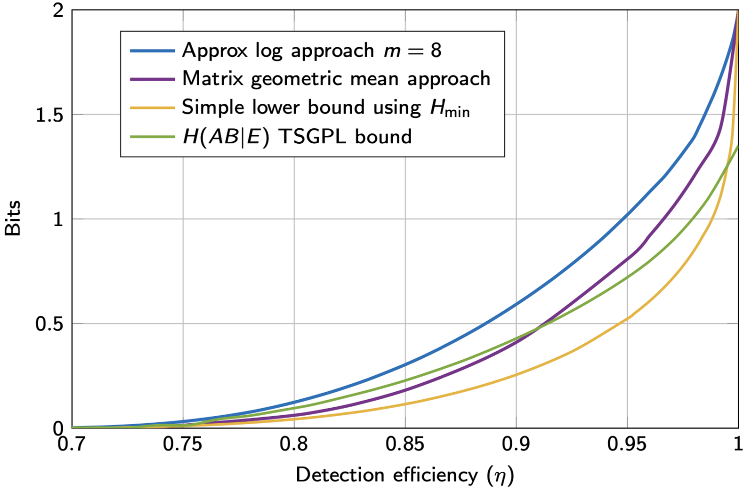
  where $H(A|E)$ is the conditional entropy.
- **Approach:** Replace $H(A|E)$ by $H_{(m)}(A|E) \leq H(A|E)$ which has a variational expression with a polynomial objective
- Lower bound on rate of the protocol:

$$\begin{array}{ll} \inf\limits_{\substack{M,N,\rho \\ a_1,\dots,a_m}} & \mathrm{tr}[\rho S(M,N,a_1,\dots,a_m)] \\ \text{s.t.} & \text{nc polynomial constraints.} \end{array}$$

- Can apply NPA hierarchy to this problem. Method we obtain is computationally faster and more accurate than previous methods



[P. Brown, H. Fawzi and O. Fawzi, Device-independent lower bounds on the conditional von Neumann entropy, arXiv:2106.13692]



[P. Brown, H. Fawzi and O. Fawzi, Device-independent lower bounds on the conditional von Neumann entropy, arXiv:2106.13692]

Conclusion

- Practical interior-point methods for quantum relative entropy cone
- Other applications of nonconvex quantum relative entropy optimization: squashed entanglement, ...

Conclusion

- Practical interior-point methods for quantum relative entropy cone
- Other applications of nonconvex quantum relative entropy optimization: squashed entanglement, ...

Thank you!

`arXiv:2106.13692` (joint with P. Brown and O. Fawzi)

`arXiv:2205.04581` (joint with J. Saunderson)