Entropic tensor scaling computes the quantum functionals

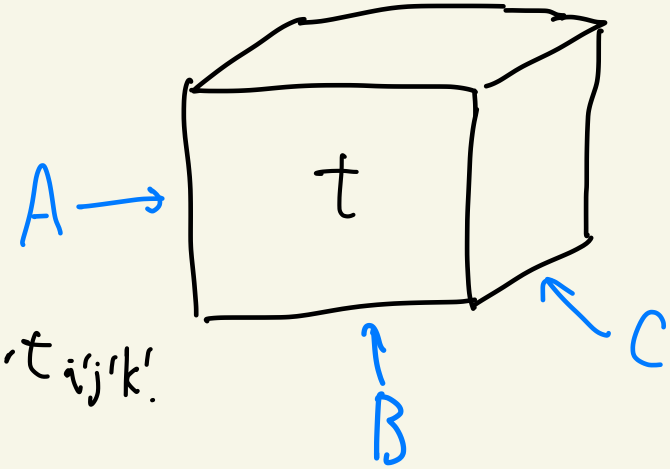
Keiya Sakabe (LMU Munich)

based on joint works with

M. Levent Doğan and Michael Walter

(3-)Tensors (generalizes to d(≥4)-tensors)

$$t = (t_{ijk}) \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3}$$



matrices

$$((A \otimes B \otimes C)\, t)_{ijk} := \sum_{i',j',k'} A_{ii'} B_{jj'} C_{kk'} t_{i'j'k'}$$

- $s \le t$: "t restricts to s."

  $\iff \exists A, B, C \quad s = (A \otimes B \otimes C)\, t$

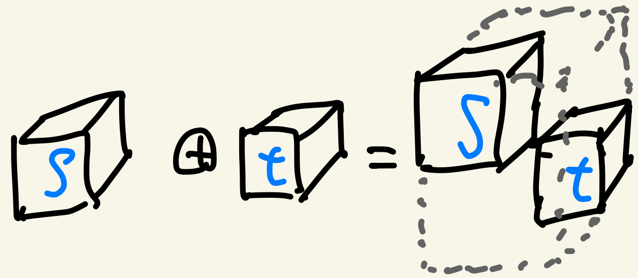
- <u>Asymptotic restriction</u>: $s \lesssim t \iff s^{\boxtimes m} \le t^{\boxtimes (m + o(m))}$
  - Captures asymptotic (sub)rank of tensors
  - Quantum resource theory

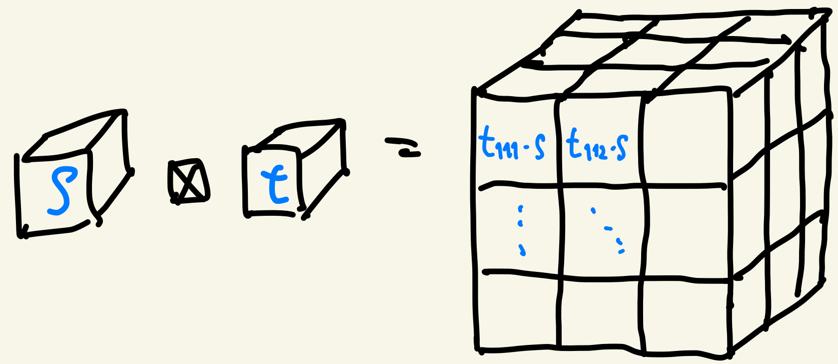
e.g. GHZ $\not\lesssim$ W. Why? We will see later.

Asymptotic spectrum of tensors [Strassen '80s]

$\mathcal{X}$: set of $\phi : \{\text{all 3-tensors}\} \to \mathbb{R}$ satisfying

$$\begin{cases} \bullet\ \phi\left(\boxed{1}\right) = 1 \\ \bullet\ \phi(s \oplus t) = \phi(s) + \phi(t) \\ \bullet\ \phi(s \boxtimes t) = \phi(s)\,\phi(t) \\ \bullet\ s \le t \Rightarrow \phi(s) \le \phi(t) \end{cases}$$





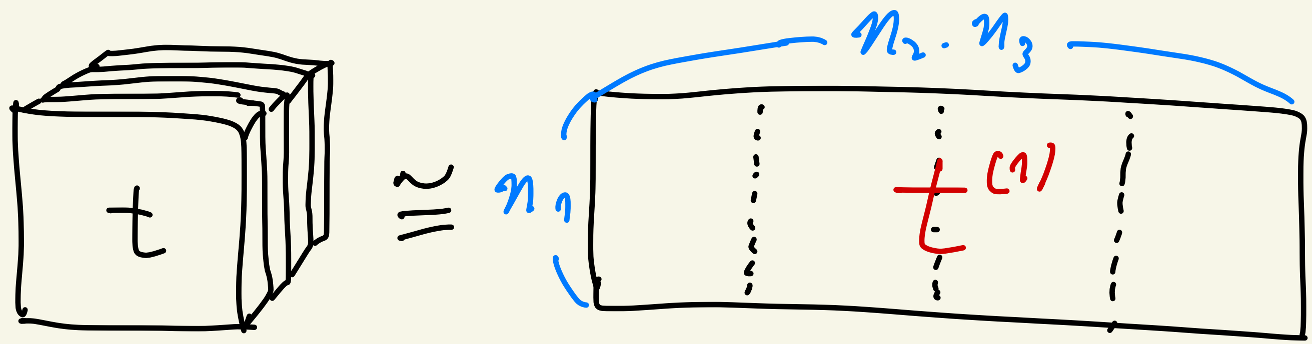
$\phi \in \mathcal{X}$: "spectral point"

Thm. [Strassen '80] . $s \lesssim t \Leftrightarrow \forall \phi \in \mathcal{X},\ \phi(s) \le \phi(t)$

Only known spectral points: Quantum functionals (next pages)

One-body marginals

Flattening:



$$\mu_1(t) := t^{(1)\dagger} t^{(1)} / \|t\|^2 \in PSD(n_1), \quad \operatorname{tr} \mu_1(t) = 1.$$

Similarly, $\mu_2(t) \in PSD(n_2)$, $\mu_3(t) \in PSD(n_3)$

In quantum information:

t: pure state on 3 particles

$\mu_1(t), \mu_2(t), \mu_3(t)$: single-partite marginals

Entanglement polytope [Walter et al. '14]

$$\Delta(t) := \{(\lambda(\mu_1(s)), \lambda(\mu_2(s)), \lambda(\mu_3(s))) \mid s \in GL \cdot t\}$$

$$(GL \cdot t := (A \otimes B \otimes C)\, t, \quad A, B, C \in GL)$$

$$\boxed{\begin{array}{l}\text{Thm. [Kirwan'84, Ness-Mumford'84, ...]} \\ \quad \Delta(t) \text{ is a polytope}\end{array}}$$

Connections

- $\Delta(\cdot)$ is GL-orbit invariant
- Quantum marginal problem: determine $x \in \Delta(t)$? t: generic
- Generalizes to other representations $\rightsquigarrow$ Moment polytope

Quantum functional [Christandl, Vrana, Zuiddam '18]

Fix $\theta = (\theta_1, \theta_2, \theta_3)$ $(\theta_i \geq 0, \ \sum_i \theta_i = 1)$

- $E_\theta(t) := \max\limits_{(x_1, x_2, x_3) \in \Delta(t)} \sum\limits_{i=1}^{3} \theta_i H(x_i)$ ← Shannon entropy

  $= \sup\limits_{s \in GL \cdot t} h_\theta(s)$ $\left(h_\theta(s) := \sum_i \theta_i H(\mu_i(s))\right)$

  ↑ von Neumann entropy

- $F_\theta(t) := 2^{E_\theta(t)}$ ← quantum functional

> **Thm** [Christandl, Vrana, Zuiddam '18]
>
> $\forall \theta$, F_θ is a spectral point (i.e., $F_\theta \in \mathcal{X}$)

- If $\theta > 0$, then $F_\theta(t) = n_1^{\theta_1} n_2^{\theta_2} n_3^{\theta_3} \iff (\mathbb{1}/n_1, \mathbb{1}/n_2, \mathbb{1}/n_3) \in \Delta(t)$.

Quantum functional computation

$$\boxed{\text{Want to compute: } E_\theta(t) = \max_{(x_1,x_2,x_3)\in\Delta(t)} \sum_i \theta_i H(x_i) = \sup_{S\in GL\cdot t} h_\theta(S)}$$

No efficient algorithm has been known!

Concave maximization over a polytope … easy???

Difficulties:

- $\Delta(t)$ is very complex.
- Computation of $\Delta(t)$ [van den Berg et al., '25]
  is only practical up to $4\times4\times4$ tensors.

Entropic tensor scaling algorithm

$$\boxed{\text{Want to compute: } E_\theta(t) = \max_{(x_1,x_2,x_3)\in\Delta(t)} \sum_i \theta_i H(x_i) = \sup_{S\in GL\cdot t} h_\theta(S)}$$

Our algorithm: Entropic Tensor Scaling

$$t_{k+1} := \left(\mu_1(t_k)^{-\frac{\theta_1}{2}} \otimes \mu_2(t_k)^{-\frac{\theta_2}{2}} \otimes \mu_3(t_k)^{-\frac{\theta_3}{2}}\right) t_k$$

Thm. [Doğan-S-Walter] $E_\theta(t_0) - h_\theta(t_k) = O(k^{-1})$.

Constant in $O(\cdot)$ is poly(input size).

- First efficient algorithm for quantum functional computation!
- When $\theta > 0$, then $\mu_i(t_k) \to I/n_i$ if $(\mathbb{1}/n_1, \mathbb{1}/n_2, \mathbb{1}/n_3) \in \Delta(t)$

  $\rightsquigarrow$ New algorithm for <u>tensor scaling problem</u>

Quantum functional: examples

$\theta = (1/3, 1/3, 1/3)$

$$\boxed{t_{k+1} := \left(\mu_1(t_k)^{-\frac{1}{6}} \otimes \mu_2(t_k)^{-\frac{1}{6}} \otimes \mu_3(t_k)^{-\frac{1}{6}}\right) t_k \quad \text{computes } E_\theta(t_0).}$$

- $t_0 := W = e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1$

$$\Rightarrow \mu_1 = \mu_2 = \mu_3 = \begin{pmatrix} 2/3 & 0 \\ 0 & 1/3 \end{pmatrix}$$

$$\Rightarrow t_1 = (2/3)^{-1/6} \cdot (2/3)^{-1/6} \cdot (1/3)^{-1/6} W \propto W,$$

W is a fixed point

$$\Rightarrow E_\theta(W) = H(2/3, 1/3).$$

- $GHZ = e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_2 \Rightarrow \mu_1 = \mu_2 = \mu_3 = I/2.$

$\Rightarrow E_\theta(GHZ) = H(1/2, 1/2)$ (entropy is maximized at uniform)

$E_\theta(GHZ) > E_\theta(W) \Rightarrow GHZ \not\leq W$!

More general convex optimization over $\Delta(t)$

$$\boxed{\text{minimize} \quad \mathcal{S}(x_1, x_2, x_3), \quad (x_1, x_2, x_3) \in \Delta(t)}$$

$\mathcal{S} : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \to \overline{\mathbb{R}} \cup \{\infty\}$ is

$$\begin{cases} \cdot \text{ lower-semicontinuous convex} \\ \cdot \ \mathcal{S}(P_1 x_1, P_2 x_2, P_3 x_3) = \mathcal{S}(x_1, x_2, x_3) \quad (P_i : \text{permutation}) \end{cases}$$

$\Rightarrow$ $\Phi : \mathrm{Herm}(n_1) \times \mathrm{Herm}(n_2) \times \mathrm{Herm}(n_3) \to \mathbb{R} \cup \{\infty\}$,

$\Phi(Y_1, Y_2, Y_3) := \mathcal{S}(\lambda(Y_1), \lambda(Y_2), \lambda(Y_3))$ is naturally defined.

Quantum functional case:

$$\mathcal{S}(x_1, x_2, x_3) = -\sum_i \theta_i H(x_i)$$

General descent algorithm

$$\boxed{\text{minimize} \quad S(x_1, x_2, x_3), \quad (x_1, x_2, x_3) \in \Delta(t)}$$

Descent algorithm for general S

$$t_{k+1} := \left(\bigotimes_{i=1}^{3} e^{-\eta \nabla_i \Phi(\mu_1(t_k), \mu_2(t_k), \mu_3(t_k))} \right) t_k$$

$(\Phi(Y_1, Y_2, Y_3) := S(\lambda(Y_1), \lambda(Y_2), \lambda(Y_3)), \ \eta > 0 : \text{step size})$

Quantum functional case:

$$S(x_1, x_2, x_3) = -\sum_i \theta_i \underset{\text{Shannon entropy}}{H}(x_i), \quad \Phi(Y_1, Y_2, Y_3) = -\sum_i \theta_i \underset{\text{von Neumann entropy}}{H}(Y_i),$$

$$e^{-\eta \nabla_i \Phi(\mu)} = e^{-\eta \theta_i (I + \log \mu_i)} = \text{const.}\ \mu_i^{-\eta \theta_i}$$

$\rightsquigarrow$ Entropic tensor scaling formula.

General descent algorithm

$$\boxed{\text{minimize } \quad S(x_1, x_2, x_3), \quad (x_1, x_2, x_3) \in \Delta(t)}$$

Descent algorithm for general S

$$t_{k+1} := \left(\bigotimes_{i=1}^{3} e^{-\eta \nabla_i \Phi(\mu_1(t_k), \mu_2(t_k), \mu_3(t_k))} \right) t_k$$

$(\Phi(Y_1, Y_2, Y_3) := S(\lambda(Y_1), \lambda(Y_2), \lambda(Y_3)),\ \eta > 0 : \text{step size})$

Thm [Doğan - S - Walter]

If S is L-smooth (i.e. $\nabla^2 S \preceq LI$),

then by taking $\eta = 1/12L$ (generally, $1/4dL$ for d-tensors),

$$\Phi(\mu_1(t_k), \mu_2(t_k), \mu_3(t_k)) - (\mathrm{Opt}) \leq O(k^{-1} \cdot \mathrm{poly}(\mathrm{input}))$$

Another application: G-stable rank

For $\alpha \in \mathbb{R}^3_{>0}$, $\mathrm{rk}^G_\alpha(t)$: G-stable rank [Derksen'22]

$$\boxed{\frac{1}{\mathrm{rk}^G_\alpha(t)} = \min_{(x_1,x_2,x_3)\in\Delta(t)} \left\|\left(\frac{x_1}{\alpha_1},\frac{x_2}{\alpha_2},\frac{x_3}{\alpha_3}\right)\right\|_\infty.}$$ [Hirai'25]

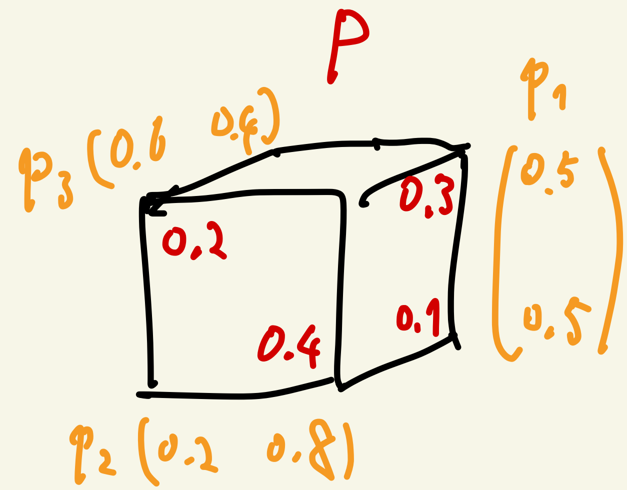
This corresponds to $S(x_1,x_2,x_3) = \left\|\left(\frac{x_1}{\alpha_1},\frac{x_2}{\alpha_2},\frac{x_3}{\alpha_3}\right)\right\|_\infty$.

 S is not smooth.

$\rightsquigarrow$ use smooth approx $\|p\|_\infty \simeq \varepsilon \log \sum_i e^{p_i/\varepsilon}$.

$\rightsquigarrow$ $\frac{1}{\mathrm{rk}^G_\alpha(t)}$ can be computed up to any predetermined error.

Minimax Formula



Support polytope

$\Omega(t) :=$

$$\{(p_1, p_2, p_3) : \text{marginal of prob. dist. } P \mid P_{ijk} > 0 \Rightarrow t_{ijk} \neq 0\}$$

For the present minimization problem (S: l.s.c. convex, symmetric),

Thm [S - Doğan - Walter]

$$\min_{(x_1,x_2,x_3) \in \Delta(t)} S(x_1, x_2, x_3) = \max_{s \in GL \cdot t} \; \min_{(p_1,p_2,p_3) \in \Omega(s)} S(p_1, p_2, p_3)$$

$\Rightarrow$ Characterized by zero/non-zero patterns obtainable by GL action!

Support functional

$$\boxed{\min_{(x_1,x_2,x_3)\in\Delta(t)} \mathcal{S}(x_1,x_2,x_3) = \max_{s\in GL\cdot t}\ \min_{(p_1,p_2,p_3)\in\Omega(s)} \mathcal{S}(p_1,p_2,p_3)}$$

In quantum functional case: $\mathcal{S}(x_1,x_2,x_3) = -\sum_i \theta_i H(x_i)$

Previously known:

- $(\mathrm{LHS}) = -\log F_\theta(t)$ ← Quantum functional
- $(\mathrm{RHS}) = -\log \zeta^\theta(t)$ ← Support functional [Strassen '80s]

We now have: $\boxed{F_\theta(t) = \zeta^\theta(t)}$ ($\forall t$: complex tensor)

<u>History</u>: Strassen conjectured that ζ^θ is a spectral point, but he could not prove it. Only $F_\theta \le \zeta^\theta$ has been known. We solved his ≥35-year-old conjecture by $F_\theta = \zeta^\theta$!

Summary

Target problem: $\boxed{\min_{(x_1,x_2,x_3)\in\Delta(t)} S(x_1,x_2,x_3)}$

S: l.s.c., convex, symmetric in each argument.

Captures
- Quantum functional ($S \sim -$entropy)
- G-stable rank ($S \sim \|\cdot\|_\infty$)

- Computable by a simple descent algorithm
- Characterized by non-zero patterns on GL-orbit

Math behind: geodesically convex optimization, generalized gradient flow [Hirai'25] (not explained today)