Asymptotic dimensional analysis on unitary representation and its application to measuring quantum relative entropy

arXiv:2406.17299

Masahito Hayashi

School of Data Science, The Chinese University of Hong Kong, Shenzhen

Shenzhen International Quantum Academy (SIQA),

Graduate School of Mathematics, Nagoya University

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

深圳国际量子研究院
Shenzhen International Quantum Academy

NAGOYA UNIVERSITY

Contents

- Asymptotic dimensional analysis on unitary representation
- Estimation of relative entropy

Schur duality

$$\mathcal{H}^{\otimes n} = \bigoplus_{\lambda \in Y_d^n} \mathcal{U}_\lambda \otimes \mathcal{V}_\lambda$$

$$\sum_{j=1}^{d} \lambda_j = n \qquad d := \dim \mathcal{H}$$

$$\lambda = (\lambda_1, \ldots, \lambda_d) \in Y_d^n$$

$$\lambda_1 \le \lambda_2 \le \ldots \le \lambda_d$$

$$d_\lambda := \dim \mathcal{U}_\lambda = \prod_{1 \le i < j \le d} \frac{j - i + \lambda_j - \lambda_i}{j - i} \le (n+1)^{\frac{d(d-1)}{2}}$$

$$d_{n,d} := \dim \bigoplus_{\lambda \in Y_d^n} \mathcal{U}_\lambda = \sum_{\lambda \in Y_d^n} d_\lambda \le (n+1)^{\frac{(d+2)(d-1)}{2}}$$

When *d* is fixed and only *n* increases,

$$\log d_{n,d} = O(\log n)$$

What happens when *d* and *n* increase?

Case with $n = O(d^{2+t})$

When $n = O(d^{2+t})$

$$\log d_\lambda = \log \prod_{1 \le i < j \le d} \frac{j - i + \lambda_j - \lambda_i}{j - i}$$

$$= \sum_{1 \le i < j \le d} \log \frac{j - i + \lambda_j - \lambda_i}{j - i} = \sum_{1 \le i < j \le d} \log\left(1 + \frac{\lambda_j - \lambda_i}{j - i}\right)$$

Sum of $O(d^2)$ terms with $\log\left(1 + \frac{\lambda_j - \lambda_i}{j - i}\right) \le \log n = O(\log d)$

It is clear that $O(d^2) \le \log d_\lambda \le O(d^2 \log d)$

Case with $n = O(d^{2+t})$

Theorem

When $n = O(d^{2+t})$,

$$\log d_{n,d} = \begin{cases} O(d^2) & \text{when } t \le 0 \\ O(d^2 \log d) & \text{when } t > 0 \end{cases}$$

Proof: $\log d_\lambda = \sum_{l=1}^{d} \sum_{i=1}^{d-l} \log(1 + \frac{\lambda_{l+i} - \lambda_i}{l})$

$$= \sum_{l=1}^{d} (d-l) \sum_{i=1}^{d-l} \frac{1}{d-l} \log(1 + \frac{\lambda_{l+i} - \lambda_i}{l})$$

$$\le \sum_{l=1}^{d} (d-l) \log(1 + \sum_{i=1}^{d-l} \frac{1}{d-l} \frac{\lambda_{l+i} - \lambda_i}{l})$$

Case with $n = cd^2$.

$$\log d_\lambda = \log \prod_{1 \le i < j \le d} \frac{j - i + \lambda_j - \lambda_i}{j - i}$$

$$\le \sum_{l=1}^{d} (d-l) \log(1 + \sum_{i=1}^{d-l} \frac{1}{d-l} \frac{\lambda_{l+i} - \lambda_i}{l})$$

$$\le \sum_{l=1}^{d} (d-l) \log(1 + \frac{1}{(d-l)} \sum_{i=1}^{d-l} \frac{\lambda_{l+i}}{l})$$

$$\le \sum_{l=1}^{d} (d-l) \log(1 + \frac{1}{(d-l)} \frac{n}{l}) = \sum_{l=1}^{d} (d-l) \log(1 + \frac{cd^2}{(d-l)l})$$

$$= d \sum_{l=1}^{d} (1 - \frac{l}{d}) \log(1 + \frac{c}{(1 - \frac{l}{d}) \frac{l}{d}})$$

Case with $n = O(d^2)$

$$\frac{1}{d^2}\log d_\lambda \le \frac{1}{d^2} d \sum_{l=1}^{d} (1-\frac{l}{d}) \log(1+\frac{c}{(1-\frac{l}{d})\frac{l}{d}})$$

$$\le \int_0^1 (1-x)\log(1+\frac{c}{(1-x)x})dx = \int_0^1 \frac{(1-x)}{s}\log(1+\frac{c}{(1-x)x})^s dx$$

$$\le \int_0^1 \frac{(1-x)}{s}\log(1+(\frac{c}{(1-x)x})^s)dx \qquad x = l/d$$

$$\le \int_0^1 \frac{(1-x)}{s}(\frac{c}{(1-x)x})^s dx \le \frac{c^s}{s(1-s)}$$

$$\log d_{n,d} = \log Y_d^n + \log \max d_\lambda$$

$$\log d_{n,d} = O(d^2) \text{ when } t \le 0 \quad n = O(d^{2+t})$$

Case with $n = O(d^{2+t}), t > 0$

We chose $\lambda_j = c' j^{1+t}$

$$\log d_\lambda = \log \prod_{1 \le i < j \le d} \frac{j - i + \lambda_j - \lambda_i}{j - i}$$

$$= \sum_{1 \le i < j \le d} \log \frac{j - i + \lambda_j - \lambda_i}{j - i} = \sum_{1 \le i < j \le d} \log(1 + \frac{\lambda_j - \lambda_i}{j - i})$$

$$= \sum_{l=1}^{d} \sum_{i=1}^{d-l} \log(1 + \frac{\lambda_{l+i} - \lambda_i}{l}) = \sum_{l=1}^{d} \sum_{i=1}^{d-l} \log(1 + \frac{c'((l+i)^{1+t} - i^{1+t})}{l})$$

$$\ge \sum_{l=d/2}^{d} \sum_{i=1}^{d-l} \log(1 + \frac{c'((l+i)^{1+t} - i^{1+t})}{l})$$

$$\ge \sum_{l=d/2}^{d} \sum_{i=1}^{d-l} \log(1 + \frac{c'(d/2)^{1+t}(1 - 1/2^{1+t})}{d})$$

$$\ge \sum_{l=d/2}^{d} \sum_{i=1}^{d-l} \log c' d^t (1 - 1/2^{1+t}) / 2^{1+t} = O(d^2 \log d)$$

Estimation of Relative entropy

Estimation of $D(\rho\|\sigma)$ with known σ
with unknown $\rho^{\otimes n}$ $D(\rho\|\sigma) := \mathrm{Tr}\rho(\log\rho - \log\sigma)$

This problem is formulated as parameter estimation with nuisance parameter in full model.

Cramer-Rao bound with nuisance parameter is

$$V(\rho\|\sigma) := \mathrm{Tr}\rho(\log\rho - \log\sigma - D(\rho\|\sigma))^2$$

This bound is attainable when model is fixed and the number n of copies increases.

Estimation of Relative entropy

Theorem

$$\mathbf{MSE} = \frac{V(\rho\|\sigma)}{n} + o\left(\frac{1}{n}\right)$$

Cramer-Rao bound with nuisance parameter is

$$V(\rho\|\sigma) := \mathrm{Tr}\,\rho(\log\rho - \log\sigma - D(\rho\|\sigma))^2$$

This bound is attainable when model is fixed and the number n of copies increases.

Schur duality

$$\mathcal{H}^{\otimes n} = \bigoplus_{\lambda \in Y_d^n} \mathcal{U}_\lambda \otimes \mathcal{V}_\lambda$$

$$\rho^{\otimes n} = \bigoplus_{\lambda \in Y_d^n} \rho_\lambda \otimes \rho_{\lambda,mix} \qquad \sigma^{\otimes n} = \bigoplus_{\lambda \in Y_d^n} \sigma_\lambda \otimes \rho_{\lambda,mix}$$

$$nD(\rho \| \sigma) = D(\rho^{\otimes n} \| \sigma^{\otimes n})$$

$$= \sum_{\lambda \in Y_d^n} \mathrm{Tr} \rho_\lambda (\log \rho_\lambda - \log \sigma_\lambda)$$

$\left\{ \left| u_{\lambda,j} \right\rangle \right\}_j$: basis of $\mathcal{U}_\lambda$ diagonalize σ_λ

$$-\log d_\lambda$$

$$\leq \mathrm{Tr} \frac{\rho_\lambda}{\mathrm{Tr}\rho_\lambda} (\log \frac{\rho_\lambda}{\mathrm{Tr}\rho_\lambda} - \log \sigma_\lambda) + \sum_j \left\langle u_{\lambda,j} \right| \frac{\rho_\lambda}{\mathrm{Tr}\rho_\lambda} \left| u_{\lambda,j} \right\rangle \log \left\langle u_{\lambda,j} \right| \sigma_\lambda \left| u_{\lambda,j} \right\rangle$$

$$\leq 0$$

Schur duality

$$\mathcal{H}^{\otimes n} = \bigoplus_{\lambda \in Y_d^n} \mathcal{U}_\lambda \otimes \mathcal{V}_\lambda$$

$$\rho^{\otimes n} = \bigoplus_{\lambda \in Y_d^n} \rho_\lambda \otimes \rho_{\lambda,mix} \qquad \sigma^{\otimes n} = \bigoplus_{\lambda \in Y_d^n} \sigma_\lambda \otimes \rho_{\lambda,mix}$$

$\left\{ |u_{\lambda,j}\rangle \right\}_j$: basis of $\mathcal{U}_\lambda$ diagonalize σ_λ

$$M_{\lambda,j} := |u_{\lambda,j}\rangle\langle u_{\lambda,j}| \otimes I(\mathcal{V}_\lambda)$$

$$x_{\lambda,j} := \frac{-1}{n} \log \mathrm{Tr} M_{\lambda,j} \sigma^{\otimes n} \qquad P^{(n)}_{\rho,\sigma}(\lambda, j) := \mathrm{Tr} M_{\lambda,j} \rho^{\otimes n}$$

$$d_{n,d} := \dim \bigoplus_{\lambda \in Y_d^n} \mathcal{U}_\lambda \le (n+1)^{\frac{(d+2)(d-1)}{2}}$$

Theorem $\mathrm{MSE}_n(\rho\|\sigma) := \sum_{(\lambda,j)} P^{(n)}_{\rho,\sigma}(\lambda, j)(x_{\lambda,j} - D(\rho\|\sigma))^2$

$$\le \left(\frac{1}{\sqrt{n}} \sqrt{V(\rho\|\sigma)} + \frac{1}{n} \log d_{n,d} \right)^2$$

When Cramer-Rao bound can be achieved?

When $n = O(d^{2+t})$,

$$\frac{1}{\sqrt{n}} \log d_{n,d} = \begin{cases} O(d^{1-t/2}) & \text{when } t \le 0 \\ O(d^{1-t/2} \log d) & \text{when } t > 0 \end{cases}$$

$$\left(\frac{1}{\sqrt{n}} \sqrt{V(\rho \| \sigma)} + \frac{1}{n} \log d_{n,d} \right)^2$$

$$= \begin{cases} \frac{1}{n} V(\rho \| \sigma) + o(\frac{1}{n}) & \text{when } t > 2 \\ O(d^{-2t} (\log d)^2) & \text{when } 0 < t \le 2 \\ O(d^{-2t}) & \text{when } 0 \ge t \end{cases}$$

Sample complexity

Assume that the minimum eigenvalue of σ_d is lower bounded by e^{-td}

We have $c_0 := \lim_{d\to\infty} \frac{1}{d^2} \max_{\rho\in S_d} V(\rho \| \sigma_d) < \infty$

Theorem

When $n = cd^2$,

$$\lim_{n\to\infty} \frac{1}{n} \log d_{n,d} \le \min_{0<s<1} \frac{c^{s-1}}{s(1-s)}$$

$$\lim_{n\to\infty} \mathrm{MSE}_n(\rho \| \sigma_d) \le \frac{(\sqrt{c_0}+4)^2}{c}$$

When $c > \frac{(\sqrt{c_0}+4)^2}{\varepsilon}$, $\lim_{n\to\infty} \mathrm{MSE}_n(\rho \| \sigma_d) \le \varepsilon$

Use of full tomography

Assume that the minimum eigenvalue of σ_d is lower bounded by e^{-td}

$$|D(\hat{\rho}\|\sigma_d) - D(\rho\|\sigma_d)| \le \|\hat{\rho} - \rho\|_1 td$$

To achieve $|D(\hat{\rho}\|\sigma_d) - D(\rho\|\sigma_d)| \le \varepsilon'$

when we employ full tomography

we need $O(d^4/\varepsilon')$ copies.

Proof:

$\|\hat{\rho} - \rho\|_1 \le \varepsilon$ requires $O(d^2/\varepsilon)$ copies.

$\|\hat{\rho} - \rho\|_1 td \le \varepsilon'$ requires $O(d^4/\varepsilon')$ copies.

Conclusion

- We have derived asymptotic behavior of the dimensions in Schur duality
- Using this, we have discussed estimation of relative entropy.
- We also derived sample complexity for upper bound for this problem.
- Our method improves conventional full tomography.

References

[1] J. Acharya, I. Issa, N. V. Shende and A. B. Wagner, "Estimating Quantum Entropy," IEEE Journal on Selected Areas in Information Theory (2020).

[2] Q. Wang and Z. Zhang, "Time-Efficient Quantum Entropy Estimator via Samplizer," to be appeared in Proceedings of the 32nd Annual European Symposium on Algorithms (ESA 2024); https://arxiv.org/abs/2401.09947 (2024).

[3] J. Haah, A. W. Harrow, Z. Ji, X.Wu and N. Yu, "Sample-Optimal Tomography of Quantum States," IEEE Transactions on Information Theory, vol. 63, no. 9, pp. 5628-5641, Sept. 2017.

[4] M. Hayashi, "Optimal sequence of quantum measurements in the sense of Stein's lemma in quantum hypothesis testing" J. Phys. A: Math. Gen., 35, 10759 – 10773 (2002).