Typicality Algorithm for Partial Trace Estimation

Robert Chen, Kevin Li, Skai Nzeuton, Yilu Pan, Yixin Wang

NYU Courant Institute of Mathematical Science

May 5th, 2023

Background Introduction

- Study of quantum systems has gained significant attention due to its potential applications in quantum computing, quantum communication, and quantum sensing
- These systems are described by density matrices, which provide a comprehensive description of the quantum state of a system
- When we have a composite quantum system described by a density matrix, we often want to study the properties of one of its subsystems without considering the rest

What is Partial Trace?

- Partial trace is a mathematical operation used to obtain the reduced density matrix from the density matrix of a larger composite system.

$$\begin{bmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{bmatrix} \xrightarrow{Tr_2} \begin{bmatrix} Tr\begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} & Tr\begin{bmatrix} a_{02} & a_{03} \\ a_{12} & a_{13} \end{bmatrix} \\ Tr\begin{bmatrix} a_{20} & a_{21} \\ a_{30} & a_{31} \end{bmatrix} & Tr\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} a_{00} + a_{11} & a_{02} + a_{13} \\ a_{20} + a_{31} & a_{22} + a_{33} \end{bmatrix}$$

Figure 1: *Partial Trace Example*

Computation and Storage Cost

- If we have an explicit representation of a matrix A, computing the partial trace of A is simple.
- What if A is a complicated matrix function?

$$Eg : A = exp(-H/t)$$

- When the size of H is $2^{20} * 2^{20}$ or even greater, the cost to compute and/or store A is too expensive.
- Hence a cheaper, faster algorithm is required.

A Stochastic Partial Trace Estimator

- Essential of setting up:

$$\mathrm{tr}(M) = \mathbb{E}\left[\mathbf{v}^T M \mathbf{v}\right], \qquad \mathrm{tr}^m(M) \approx \frac{1}{m}\sum_{i=1}^{m} \mathbf{v}^T M \mathbf{v}$$

  where v is a random vector with independent and identically distributed entries having a standard normal distribution.

- Extend this property to partial trace:

$$\begin{bmatrix} \mathrm{tr}(A_{1,1}) & \mathrm{tr}(A_{1,2}) & \cdots & \mathrm{tr}(A_{1,a}) \\ \mathrm{tr}(A_{2,1}) & \mathrm{tr}(A_{2,2}) & \cdots & \mathrm{tr}(A_{2,a}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{tr}(A_{a,1}) & \mathrm{tr}(A_{a,2}) & \cdots & \mathrm{tr}(A_{a,a}) \end{bmatrix} = \mathrm{E} \begin{bmatrix} \mathbf{v}^T A_{1,1} \mathbf{v} & \mathbf{v}^T A_{1,2} \mathbf{v} & \cdots & \mathbf{v}^T A_{1,a} \mathbf{v} \\ \mathbf{v}^T A_{2,1} \mathbf{v} & \mathbf{v}^T A_{2,2} \mathbf{v} & \cdots & \mathbf{v}^T A_{2,a} \mathbf{v} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}^T A_{a,1} \mathbf{v} & \mathbf{v}^T A_{a,2} \mathbf{v} & \cdots & \mathbf{v}^T A_{a,a} \mathbf{v} \end{bmatrix}$$

- However, introducing randomness will definitely lead to a potentially large variance of the estimator.

A Variance Reduce Algorithm

- Notice the following property of partial trace:

$$\mathrm{tr}_b(A) = \mathrm{tr}_b(\widetilde{A}) + \mathrm{tr}_b(A - \widetilde{A}).$$

- We can compute the trace of $\widetilde{A}$ explicitly, and use the stochastic partial trace estimator we developed earlier to estimate $\mathrm{tr}_b(A - \widetilde{A})$, ie:

$$\mathrm{tr}_b(\mathbf{A}) \approx \mathrm{tr}_b(\widetilde{A}) + \mathrm{tr}_b^m(A - \widetilde{A}).$$

- $\widetilde{A}$ should be chosen so that $\mathrm{tr}_b^m(A - \widetilde{A})$ produces a relatively smaller variance.
- Generalize $A = exp(-H/t)$ to any arbitrary matrix function $A = f(H)$ using Krylov Subspace Method, which uses matrix polynomial functions to approximate A.

Numerical Experiments

$$\rho^*(\beta) := \frac{tr_b(exp(-\beta H))}{tr(tr_b(exp(-\beta H)))}, \beta = \frac{1}{t}$$

$$\tilde{A} := QQ^T A Q Q^T, Q \in \mathbb{R}^{d_A * k}$$

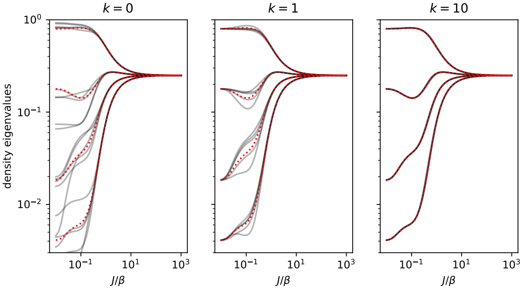
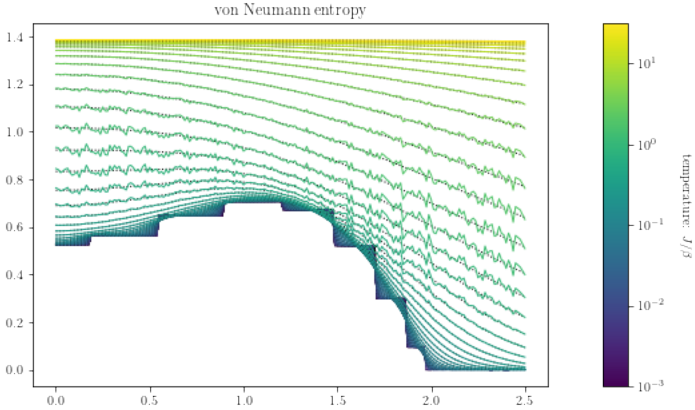


Figure 2: *comparison of $k = 0$, $k = 1$, and $k = 10$*

Adaptive Point Estimation

- Our blackbox function generates the graph below:



- To estimate the points at which the function changes, use binary search method with sample points.

Supercomputer

- To study the performance of different quantum systems, we will conduct numerical experiments on the NYU Greene Supercomputer.
- The problem is well-suited for parallel computing:
- Inputs are independent of each other without the need to exchange information.

Q & A

Any Questions?