A Typicality-Based Algorithm for Partial Trace Estimation

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

Application: Quantum Equilibrium Thermodynamics

Total system (with Hamiltonian $\mathbf{H}$) in thermal equilibrium at inverse temperature β (due to weak coupling with a "superbath") has density matrix:

$$\boldsymbol{\rho}_{\mathrm{t}}(\beta) = \frac{\exp(-\beta\mathbf{H})}{Z_{\mathrm{t}}(\beta)}, \qquad Z_{\mathrm{t}}(\beta) = \operatorname{tr}(\exp(-\beta\mathbf{H})). \tag{1}$$

The subsystem of interest has density matrix:

$$\boldsymbol{\rho}^*(\beta) = \operatorname{tr}_{\mathrm{b}}(\boldsymbol{\rho}_{\mathrm{t}}(\beta)). \tag{2}$$

Partial Traces

Let d_{s} and d_{b} be the dimension of $\mathcal{H}_{\mathrm{s}}$ and $\mathcal{H}_{\mathrm{b}}$ respectively, so that $d_{\mathrm{t}} = d_{\mathrm{s}}d_{\mathrm{b}}$ is the dimension of $\mathcal{H}_{\mathrm{t}} = \mathcal{H}_{\mathrm{s}} \otimes \mathcal{H}_{\mathrm{b}}$. A generic matrix $\mathbf{A} : \mathcal{H}_{\mathrm{t}} \to \mathcal{H}_{\mathrm{t}}$ can be partitioned as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \cdots & \mathbf{A}_{1,d_{\mathrm{s}}} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \cdots & \mathbf{A}_{2,d_{\mathrm{s}}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{d_{\mathrm{s}},1} & \mathbf{A}_{d_{\mathrm{s}},2} & \cdots & \mathbf{A}_{d_{\mathrm{s}},d_{\mathrm{s}}} \end{bmatrix}, \tag{3}$$

where $\mathbf{A}_{i,j} : \mathcal{H}_{\mathrm{b}} \to \mathcal{H}_{\mathrm{b}}$ for each i, j. The partial trace of $\mathbf{A}$ over $\mathcal{H}_{\mathrm{b}}$ is defined as

$$\operatorname{tr}_{\mathrm{b}}(\mathbf{A}) := \begin{bmatrix} \operatorname{tr}(\mathbf{A}_{1,1}) & \operatorname{tr}(\mathbf{A}_{1,2}) & \cdots & \operatorname{tr}(\mathbf{A}_{1,d_{\mathrm{s}}}) \\ \operatorname{tr}(\mathbf{A}_{2,1}) & \operatorname{tr}(\mathbf{A}_{2,2}) & \cdots & \operatorname{tr}(\mathbf{A}_{2,d_{\mathrm{s}}}) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{tr}(\mathbf{A}_{d_{\mathrm{s}},1}) & \operatorname{tr}(\mathbf{A}_{d_{\mathrm{s}},2}) & \cdots & \operatorname{tr}(\mathbf{A}_{d_{\mathrm{s}},d_{\mathrm{s}}}) \end{bmatrix}. \tag{4}$$

Stochastic (Partial) Trace Estimation

Suppose $\mathbf{v}$ has independent and identically distributed standard normal entries. Then,

$$\mathbb{E}\big[\mathbf{v}^{\mathsf{T}}\mathbf{M}\mathbf{v}\big] = \operatorname{tr}(\mathbf{M}), \qquad \mathbb{V}\big[\mathbf{v}^{\mathsf{T}}\mathbf{M}\mathbf{v}\big] = 2\|\mathbf{M}\|_{\mathsf{F}}^2.$$

Such estimators can be extended to partial traces [1]. In particular,

$$(\mathbf{I}_{d_{\mathrm{s}}} \otimes \mathbf{v})^{\mathsf{T}}\mathbf{A}(\mathbf{I}_{d_{\mathrm{s}}} \otimes \mathbf{v}) = \begin{bmatrix} \mathbf{v}^{\mathsf{T}}\mathbf{A}_{1,1}\mathbf{v} & \mathbf{v}^{\mathsf{T}}\mathbf{A}_{1,2}\mathbf{v} & \cdots & \mathbf{v}^{\mathsf{T}}\mathbf{A}_{1,d_{\mathrm{s}}}\mathbf{v} \\ \mathbf{v}^{\mathsf{T}}\mathbf{A}_{2,1}\mathbf{v} & \mathbf{v}^{\mathsf{T}}\mathbf{A}_{2,2}\mathbf{v} & \cdots & \mathbf{v}^{\mathsf{T}}\mathbf{A}_{2,d_{\mathrm{s}}}\mathbf{v} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}^{\mathsf{T}}\mathbf{A}_{d_{\mathrm{s}},1}\mathbf{v} & \mathbf{v}^{\mathsf{T}}\mathbf{A}_{d_{\mathrm{s}},2}\mathbf{v} & \cdots & \mathbf{v}^{\mathsf{T}}\mathbf{A}_{d_{\mathrm{s}},d_{\mathrm{s}}}\mathbf{v} \end{bmatrix} \tag{5}$$

Given independent and identically distributed copies $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of $\mathbf{v}$, we arrive at an estimator

$$\widehat{\operatorname{tr}}_{\mathrm{b}}^m(\mathbf{A}) := \frac{1}{m}\sum_{i=1}^{m}(\mathbf{I}_{d_{\mathrm{s}}} \otimes \mathbf{v}_i)^{\mathsf{T}}\mathbf{A}(\mathbf{I}_{d_{\mathrm{s}}} \otimes \mathbf{v}_i). \tag{6}$$

The variance of a random matrix $\mathbf{X}$ can be defined as $\mathbb{V}[\mathbf{X}] := \mathbb{E}\Big[\big\|\mathbf{X} - \mathbb{E}[\mathbf{X}]\big\|_{\mathsf{F}}^2\Big]$. For our estimator:

$$\mathbb{V}\Big[\widehat{\operatorname{tr}}_{\mathrm{b}}^m(\mathbf{A})\Big] = \frac{1}{m}\mathbb{V}\Big[(\mathbf{I}_{d_{\mathrm{s}}} \otimes \mathbf{v})^{\mathsf{T}}\mathbf{A}(\mathbf{I}_{d_{\mathrm{s}}} \otimes \mathbf{v})\Big] = \frac{1}{m}\sum_{i,j=1}^{d_{\mathrm{s}}} 2\|\mathbf{A}_{i,j}\|_{\mathsf{F}}^2 = \frac{2}{m}\|\mathbf{A}\|_{\mathsf{F}}^2. \tag{7}$$

A Variance-Reduced Algorithm

The linearity of partial trace implies that, for any matrix $\widetilde{\mathbf{A}}$,

$$\operatorname{tr}_{\mathrm{b}}(\mathbf{A}) = \operatorname{tr}_{\mathrm{b}}(\widetilde{\mathbf{A}}) + \operatorname{tr}_{\mathrm{b}}(\mathbf{A} - \widetilde{\mathbf{A}}). \tag{8}$$

We can then estimate the partial trace of $\operatorname{tr}_{\mathrm{b}}(\mathbf{A})$ by computing the partial trace of the first term exactly, and applying the randomized estimator (5) to the residual term:

$$\operatorname{tr}_{\mathrm{b}}(\mathbf{A}) \approx \operatorname{tr}_{\mathrm{b}}(\widetilde{\mathbf{A}}) + \widehat{\operatorname{tr}}_{\mathrm{b}}^m(\mathbf{A} - \widetilde{\mathbf{A}}). \tag{9}$$

The variance of such an estimate is entirely due to the variance of $\widehat{\operatorname{tr}}_{\mathrm{b}}^m(\mathbf{A} - \widetilde{\mathbf{A}})$. If $\|\mathbf{A} - \widetilde{\mathbf{A}}\|_{\mathsf{F}}^2 < \|\mathbf{A}\|_{\mathsf{F}}^2$, then the variance of the estimator on the right of (9) is reduced over that of $\widehat{\operatorname{tr}}_{\mathrm{b}}^m(\mathbf{A})$.

Splittings similar to (9) have previously been used as a variance reduction technique for regular trace estimation [2, 3, etc.].

Using the fact that $\operatorname{tr}(\mathbf{x}_{(i)}\mathbf{x}_{(j)}^{\mathsf{T}}) = \mathbf{x}_{(i)}^{\mathsf{T}}\mathbf{x}_{(j)}$, we find that

$$\operatorname{tr}_{\mathrm{b}}(\mathbf{x}\mathbf{x}^{\mathsf{T}}) = \begin{bmatrix} \mathbf{x}_{(1)}^{\mathsf{T}}\mathbf{x}_{(1)} & \mathbf{x}_{(2)}^{\mathsf{T}}\mathbf{x}_{(1)} & \cdots & \mathbf{x}_{(d_{\mathrm{s}})}^{\mathsf{T}}\mathbf{x}_{(1)} \\ \mathbf{x}_{(1)}^{\mathsf{T}}\mathbf{x}_{(2)} & \mathbf{x}_{(2)}^{\mathsf{T}}\mathbf{x}_{(2)} & \cdots & \mathbf{x}_{(d_{\mathrm{s}})}^{\mathsf{T}}\mathbf{x}_{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{(1)}^{\mathsf{T}}\mathbf{x}_{(d_{\mathrm{s}})} & \mathbf{x}_{(2)}^{\mathsf{T}}\mathbf{x}_{(d_{\mathrm{s}})} & \cdots & \mathbf{x}_{(d_{\mathrm{s}})}^{\mathsf{T}}\mathbf{x}_{(d_{\mathrm{s}})} \end{bmatrix}. \tag{10}$$

This allows us to take $\widetilde{\mathbf{A}} = \mathbf{Q}\mathbf{Q}^{\mathsf{T}}\mathbf{A}\mathbf{Q}\mathbf{Q}^{\mathsf{T}} = \sum_{i=1}^{k}\theta_i\mathbf{x}_i\mathbf{x}_i^{\mathsf{T}}$ as a low-rank approximation of $\mathbf{A}$ which gives the approximation,

$$\widehat{\operatorname{tr}}_{\mathrm{b}}^m(\mathbf{A}; \mathbf{Q}) := \sum_{i=1}^{k}\theta_i\operatorname{tr}_{\mathrm{b}}(\mathbf{x}_i\mathbf{x}_i^{\mathsf{T}}) + \widehat{\operatorname{tr}}_{\mathrm{b}}^m(\mathbf{A} - \mathbf{Q}\mathbf{Q}^{\mathsf{T}}\mathbf{A}\mathbf{Q}\mathbf{Q}^{\mathsf{T}}) \tag{11}$$

We take $\mathbf{Q}$ as the top k eigenvectors of $\mathbf{A}$, which minimizes the variance of the algorithm and helps with numerical stability in the case that products with $\mathbf{A}$ are inexact, such as if products with $f(\mathbf{H})$ are approximated using the Lanczos algorithm.

In this case, we have:

$$\|\mathbf{A} - \mathbf{Q}\mathbf{Q}^{\mathsf{T}}\mathbf{A}\mathbf{Q}\mathbf{Q}^{\mathsf{T}}\|_{\mathsf{F}}^2 = \min_{\operatorname{rank}(\widetilde{\mathbf{A}})=k}\|\mathbf{A} - \widetilde{\mathbf{A}}\|_{\mathsf{F}}^2 = \sum_{i=k+1}^{d_{\mathrm{t}}}\sigma_i^2, \tag{12}$$

where $\{\sigma_i\}$ are the singular values of $\mathbf{A}$ in non-increasing order.

Numerical Experiments

Our experiments focus on Heisenberg spin systems in an isotropic magnetic field oriented with the z-axis

$$\mathbf{H} := \sum_{i,j=1}^{N}\left[J_{i,j}^{\mathsf{x}}\boldsymbol{\sigma}_i^{\mathsf{x}}\boldsymbol{\sigma}_j^{\mathsf{x}} + J_{i,j}^{\mathsf{y}}\boldsymbol{\sigma}_i^{\mathsf{y}}\boldsymbol{\sigma}_j^{\mathsf{y}} + J_{i,j}^{\mathsf{z}}\boldsymbol{\sigma}_i^{\mathsf{z}}\boldsymbol{\sigma}_j^{\mathsf{z}}\right] + \frac{h}{2}\sum_{i=1}^{N}\boldsymbol{\sigma}_i^{\mathsf{z}}. \tag{13}$$

where $\boldsymbol{\sigma}_i^{\mathsf{x/y/z}} = \mathbf{I}_{i-1} \otimes \boldsymbol{\sigma}^{\mathsf{x/y/z}} \otimes \mathbf{I}_{N-i}$ and $\boldsymbol{\sigma}^{\mathsf{x/y/z}}$ are the Pauli matrices

Numerical Experiments cont.

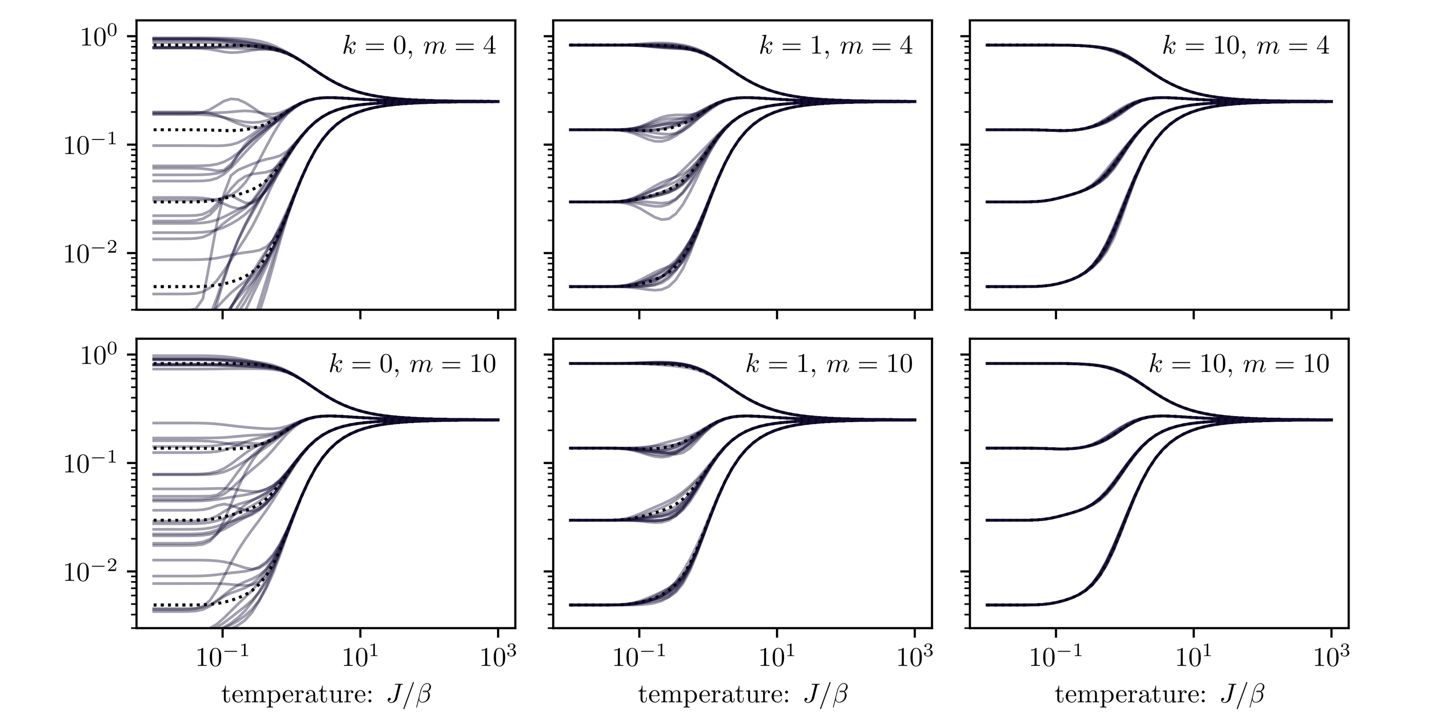

Figure 1. Comparison eigenvalues of $\boldsymbol{\rho}^*(\beta)$ when $k = 0$ (equivalent to [1]), $k = 1$, and $k = 10$. Solid lines correspond to repeated runs of the algorithm and dotted lines the true values.

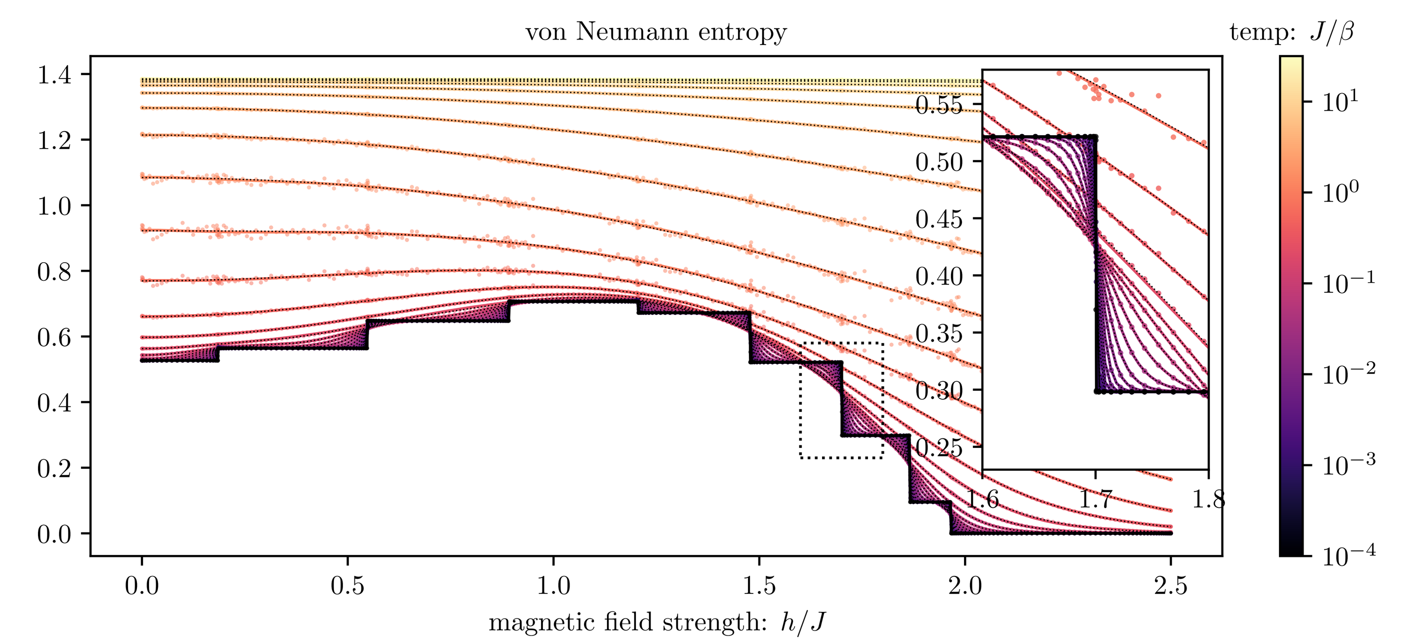

Figure 2. von Neumann Entropy $-\operatorname{tr}(\boldsymbol{\rho}^*(\beta)\ln\boldsymbol{\rho}^*(\beta))$as a function of magnetic field strength at various temperatures on the solvable model at N = 16.

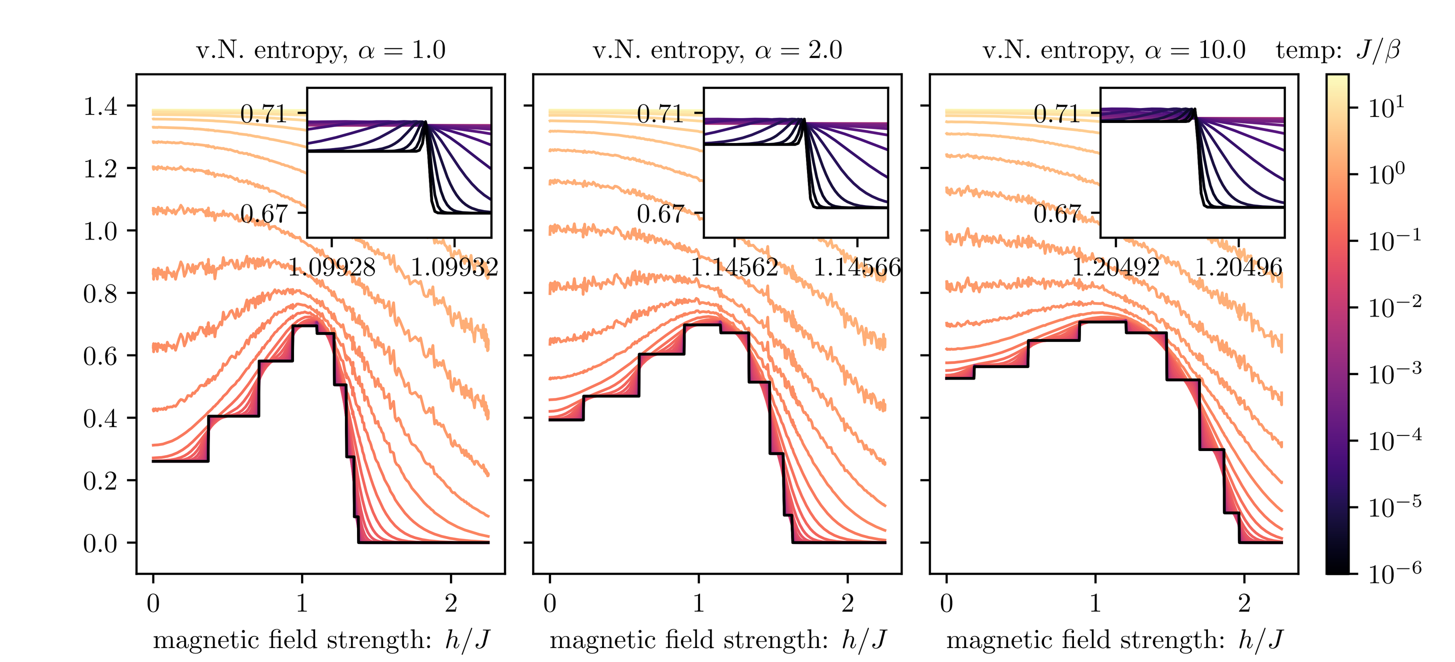

Figure 3. Von Neumann entropy for the long-range spin chain at varying values of α, h, and β. The inset figure shows the transition to the right of the top "plateau".

[1] T. Chen and Y-C. Cheng. Numerical computation of the equilibrium-reduced density matrix for strongly coupled open quantum systems. J. Chem. Phys., 157(6):064106, August 2022. [2] D. Girard. Un algorithme rapide pour le calcul de la trace de l'inverse d'une grande matrice, March 1987. [3] L. Wu, et al. Estimating the trace of the matrix inverse by interpolating from the diagonal of an approximate inverse. J. Comp. Phys., 326:828–844, 2016