Statistical Estimation in the Spiked Tensor Model via the Quantum Approximate Optimization Algorithm

Leo Zhou

NeurIPS 2024

NEURAL INFORMATION PROCESSING SYSTEMS

Caltech UCLA

Berkeley UNIVERSITY OF CALIFORNIA

Joao Basso

Song Mei

Spiked Tensor: a statistical model for learning tensor data

Given $$\boldsymbol{Y} = \frac{\lambda_n}{n^{q/2}}\boldsymbol{u}^{\otimes q} + \frac{1}{\sqrt{n}}\boldsymbol{W} \quad \in (\mathbb{R}^n)^{\otimes q}$$

Signal $\boldsymbol{u} \sim \mathrm{Unif}(\{\pm 1\}^n)$ *Noise* $W_{i_1,\ldots,i_q} \sim \mathcal{N}(0,1)$

Goal: (weak) recovery of the hidden signal

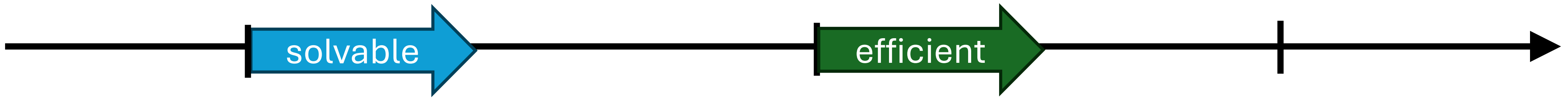
Find $\hat{\boldsymbol{u}}(\boldsymbol{Y})$ *such that* $\lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n^2}\langle \hat{\boldsymbol{u}}(\boldsymbol{Y}), \boldsymbol{u}\rangle^2\right] > 0$

Spiked Tensor problem has a huge **statistical-computational gap!**

$$\boldsymbol{Y} = \frac{\lambda_n}{n^{q/2}} \boldsymbol{u}^{\otimes q} + \frac{1}{\sqrt{n}} \boldsymbol{W}$$

Signal-to-Noise Ratio (λ_n) Thresholds for Recovery

$\Theta(1)$ $\Theta\left(n^{(q-2)/4}\right)$ $\Theta\left(n^{(q-2)/2}\right)$



λ_n

Information-theoretic threshold

Classical Computational Threshold

(achieved e.g. by spectral methods)

Tensor power iteration

Spiked Tensor problem has a huge **statistical-computational gap!**

$$\boldsymbol{Y} = \frac{\lambda_n}{n^{q/2}} \boldsymbol{u}^{\otimes q} + \frac{1}{\sqrt{n}} \boldsymbol{W}$$

Signal-to-Noise Ratio (λ_n) Thresholds for Recovery

$\Theta(1)$ | ***Statistical-Computational Gap!*** | $\Theta\left(n^{(q-2)/4}\right)$ | $\Theta\left(n^{(q-2)/2}\right)$

Information-theoretic threshold

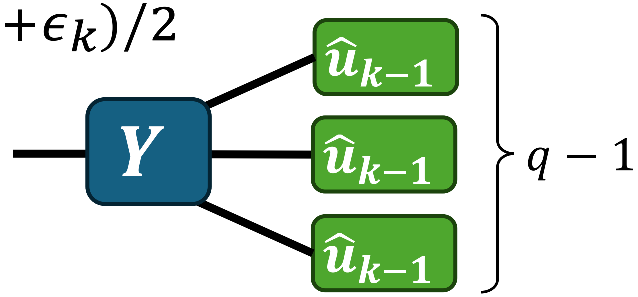
Classical Computational Threshold

(achieved e.g. by spectral methods)

Tensor power iteration

Classical algorithms for spiked tensor

- **Tensor power iteration** (k-step) works if $\lambda_n \gtrsim n^{(q-2+\epsilon_k)/2}$



$$\hat{\boldsymbol{u}}_k \propto \boldsymbol{Y}[\boldsymbol{u}_{k-1}^{\otimes q-1}] \quad \text{with} \quad \hat{\boldsymbol{u}}_0 \sim \mathrm{Unif}(\{\pm 1\}^n)$$

- **Spectral method** (with tensor unfolding) works if $\lambda_n \gtrsim n^{(q-2)/4}$

$$\mathrm{mat}(\boldsymbol{Y}) = n^{\lfloor q/2 \rfloor} \times n^{\lceil q/2 \rceil} \text{ matrix}$$

Find largest singular vector of $\mathrm{mat}(\boldsymbol{Y})$, which will have nontrivial overlap with $\mathrm{vec}(\boldsymbol{u}^{\otimes \lceil q/2 \rceil})$

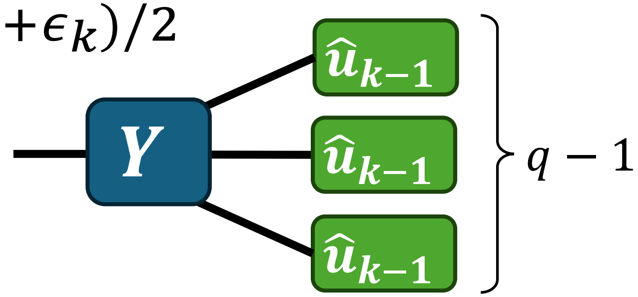
- **Maximum Likelihood Estimator** works when $\lambda_n \gtrsim \sqrt{q \log q}$

$$\hat{\boldsymbol{u}}_{\mathrm{MLE}} = \arg\max_{\boldsymbol{z}} C_{\boldsymbol{Y}}(\boldsymbol{z}) \qquad C_{\boldsymbol{Y}}(\boldsymbol{z}) = \langle \boldsymbol{Y}, \boldsymbol{z}^{\otimes q} \rangle = \sum_{i_1,\dots,i_q} Y_{i_1,\dots,i_q} z_{i_1} \cdots z_{i_q}$$

MLE is **inefficient** to compute!! NP-hard in general to find exactly

Classical algorithms for spiked tensor

- **Tensor power iteration** (k-step) works if $\lambda_n \gtrsim n^{(q-2+\epsilon_k)/2}$



$$\hat{u}_k \propto Y[u^{\otimes q-1}] \quad \text{with} \quad \hat{u}_0 \sim \mathrm{Unif}(\{\pm 1\}^n)$$

Can quantum algorithms overcome the gap?

- orks if $\lambda_n \gtrsim n^{(q-2)/4}$

Find largest singular vector of mat$(\boldsymbol{Y})$, which will have nontrivial overlap with vec$\left(\boldsymbol{u}^{\otimes \lceil q/2 \rceil}\right)$

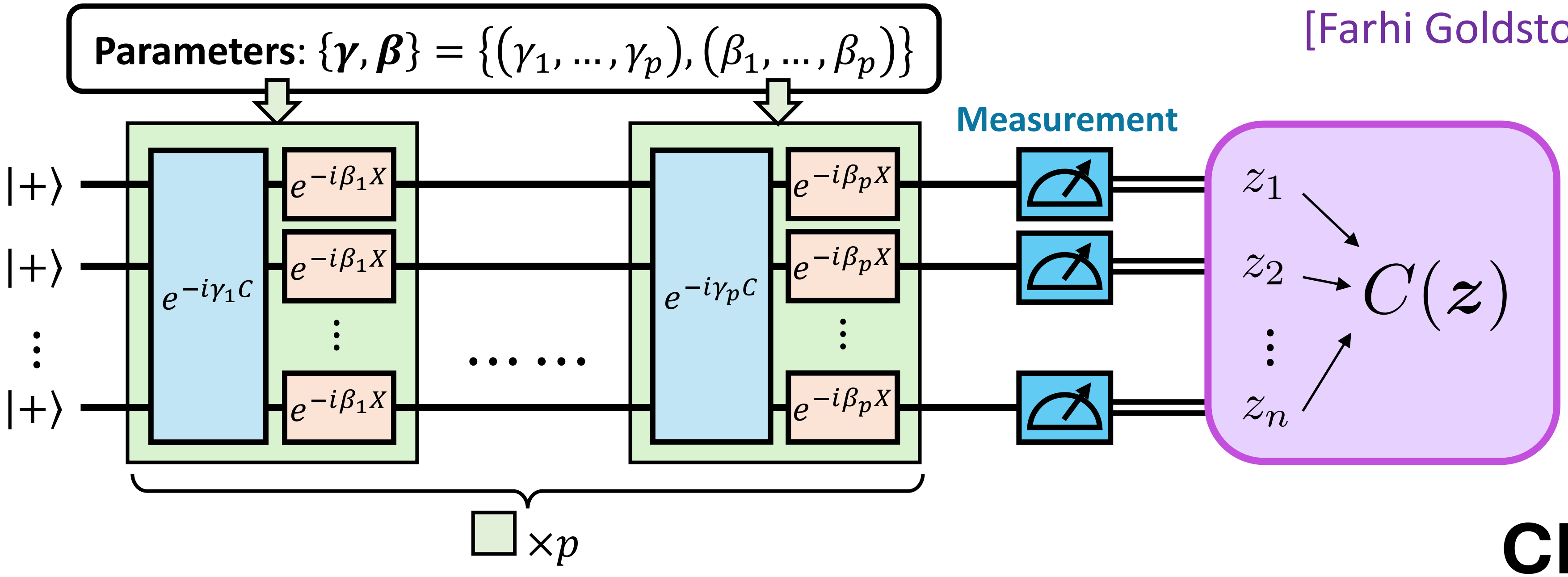
- en $\lambda_n \gtrsim \sqrt{q \log q}$

$$\hat{u}_{\mathrm{MLE}} = \arg\max_{z} C_{Y}(z) \qquad C_Y(z) = \langle Y, z^{\otimes q} \rangle = \sum_{i_1,\ldots,i_q} Y_{i_1,\ldots,i_q} z_{i_1} \cdots z_{i_q}$$

MLE is **inefficient** to compute!! NP-hard in general to find exactly

Quantum Approximate Optimization Algorithm (QAOA)

[Farhi Goldstone Gutmann 2014]



$$C\,|\boldsymbol{z}\rangle = C(\boldsymbol{z})\,|\boldsymbol{z}\rangle$$

$$B = \sum_{i=1}^{n} X_i$$

$$|\gamma, \beta\rangle = e^{-i\beta_p B} e^{-i\gamma_p C} \ldots e^{-i\beta_1 B} e^{-i\gamma_1 C} \, |+\rangle^{\otimes n}$$

Choose γ, β to maximize $\langle C \rangle$

- Simple & easy implementation (e.g. ions (2019), superconductors (2020) and atoms (2022))
- Cannot classically sample (“supremacy”) even @ p=1 [Farhi Harrow ’16] [Krovi ’22]
- ***Guaranteed to get $C_{\max}$ as depth $p \to \infty$!***

Our Result: QAOA for Spiked Tensor

- Apply QAOA to approximately maximize $C_{\boldsymbol{Y}}(\boldsymbol{z}) = \langle \boldsymbol{Y}, \boldsymbol{z}^{\otimes q} \rangle$

- Suppose $\lim_{n\to\infty} \frac{\lambda_n}{n^{(q-2+\varepsilon_p)/2}} = \Lambda$ $\qquad \varepsilon_p = \begin{cases} \frac{q-2}{(q-1)^p-1}, & q > 2, \\ 1/p, & q = 2. \end{cases}$

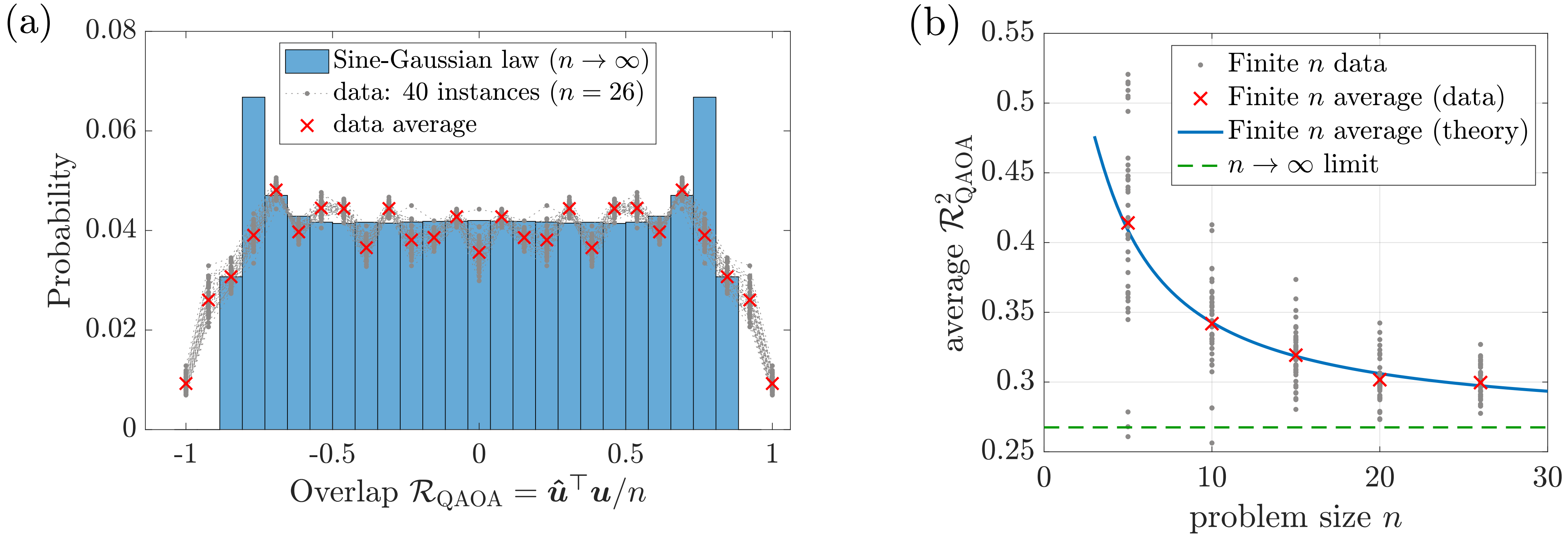
- Let $\hat{\boldsymbol{u}} \sim |\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle$ and $R_{QAOA} = \hat{\boldsymbol{u}}^T \boldsymbol{u}/n$ be its overlap with the signal, then

$$\mathcal{R}_{\mathrm{QAOA}} \xrightarrow{d} a_p \sin\left[b_p \Lambda^{1/\varepsilon_p} G^{(q-1)^p}\right] \qquad G \sim \mathcal{N}(0,1)$$

$a_p(\boldsymbol{\gamma}, \boldsymbol{\beta})$, $b_p(\boldsymbol{\gamma}, \boldsymbol{\beta})$ can be calculated classically in $\tilde{O}(4^p)$ time

QAOA for Spiked Tensor: a Sine-Gaussian Law

Example numerical simulation at ***p*=1, *q*=2**
(1-step QAOA for spiked matrix)



Comparing QAOA and Power Iteration

Let $$\lim_{n\to\infty} \frac{\lambda_n}{n^{(q-2+\varepsilon_p)/2}} = \Lambda$$

$$\varepsilon_p = \begin{cases} \frac{q-2}{(q-1)^p-1}, & q > 2, \\ 1/p, & q = 2. \end{cases}$$

Note:
$1/\epsilon_p \sim \exp(p)$
when $q > 2$

p-step QAOA

$$\mathcal{R}_{\text{QAOA}} \xrightarrow{d} a_p \sin\left[b_p \Lambda^{1/\varepsilon_p} G^{(q-1)^p}\right]$$

$$\approx (|a_p b_p|^{\varepsilon_p} \Lambda)^{1/\varepsilon_p} G^{(q-1)^p}$$

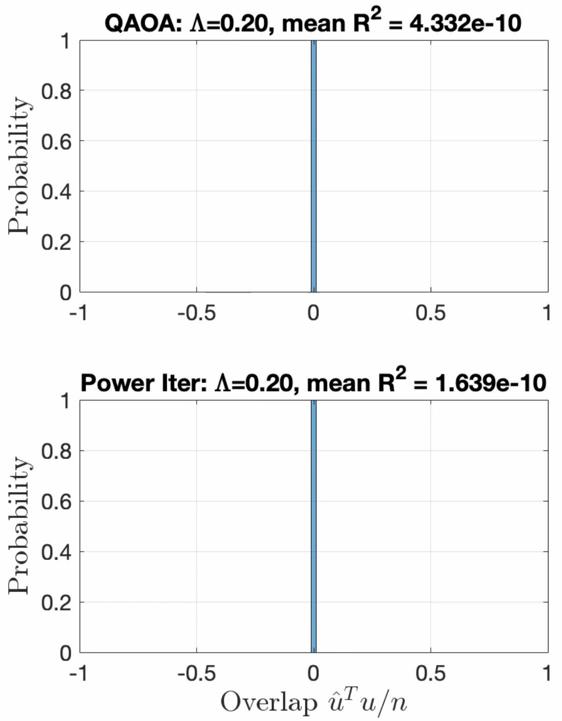
$G \sim \mathcal{N}(0,1)$

quantum enhancement factor

p-step Power Iteration

$$\mathcal{R}_{\text{PI}} \xrightarrow{d} \sin\left[\tan^{-1}\left(\Lambda^{1/\varepsilon_p} G^{(q-1)^p}\right)\right]$$

$$\approx \Lambda^{1/\varepsilon_p} G^{(q-1)^p}$$



QAOA effectively improves the recovery threshold vs. power iteration

Quantum enhancement factor $|a_p b_p|^{\varepsilon_p}$

p \ q	2	3	4	5	6	7
1	0.8578	1.0505	1.2131	1.3562	1.4857	1.6047
2	0.9663	1.0505	1.1916	1.2882	1.4167	1.5162
3	1.0204	1.0314	1.1615	1.2555	1.3844	1.4917
4	1.0487	1.0144	1.1419	1.2447	1.3795	1.4858
5	1.0631	1.0063	1.1327	1.2411	1.3770	1.4845
6	1.0697	1.0013	1.1297	1.2399	1.3743	1.4842
7	1.0719					

Effective recovery thresholds:

p-step QAOA: $\lambda_n \simeq |a_p b_p|^{-\epsilon_p} n^{(q-2+\epsilon_p)/2}$

p-step Power iteration: $\lambda_n \simeq n^{(q-2+\epsilon_p)/2}$

Classical computational threshold: $\lambda_n \propto n^{(q-2)/4}$

p-step QAOA with tensor unfolding: $\lambda_n \simeq |a_p b_p|^{-\epsilon_p} n^{(q-2+q/p)/4}$

QAOA effectively improves the recovery threshold vs. power iteration

Quantum enhancement factor

$|a_p b_p|^{\varepsilon_p}$

p \ q	2	3	4	5	6	7
1	0.8578	1.0505	1.2131	1.3562	1.4857	1.6047
2	0.9663	1.0505	1.1916	1.2882	1.4167	1.5162
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.4917
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.4858
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.4845
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.4842

For more details, see

Poster # 94828

Wednesday 11am – 2pm

December 11

Effective recovery thresholds:

Classical computational threshold: $\lambda_n \propto n^{(q-2)/4}$

p-step QAOA with tensor unfolding: $\lambda_n \simeq |a_p b_p|^{-\epsilon_p} n^{(q-2+q/p)/4}$