Rank Is All You Need:
Estimating the Trace of Powers of Density Matrices

Junseo Lee

Collaborators: Myeongjin Shin, Seungwoo Lee, Kabgyun Jeong

arXiv:2408.00314

"____ Is all you need" papers …

Attention Is All You Need

+ 135180 citations

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

"____ Is all you need" papers …

Attention Is All You Need

+ 135180 citations

Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.com
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

and many others …!

Awesome *"all you need"* papers

Last Added 2024-09-26 | Update Paper List failing

About

This repository is a list of all "all you need" papers. The list is updated daily using the arXiv API.

Paper List

Title	Authors	Date
CNN Is All You Need	Qiming Chen et al.	2017-12-27
Bytes are All You Need: End-to-End Multilingual Speech Recognition and Synthesis with Bytes	Bo Li et al.	2018-11-22
Depthwise Convolution is All You Need for Learning Multiple Visual Domains	Yunhui Guo et al.	2019-02-03
CBOW Is Not All You Need: Combining CBOW with the Compositional Matrix Space Model	Florian Mai et al.	2019-02-18
All You Need is a Few Shifts: Designing Efficient Convolutional Neural Networks for Image Classification	Weijie Chen et al.	2019-03-13
ThumbNet: One Thumbnail Image Contains All You Need for Recognition	Chen Zhao et al.	2019-04-10
Focus Is All You Need: Loss Functions For Event-based Vision	Guillermo Gallego et al.	2019-04-15

https://github.com/KentoNishi/awesome-all-you-need-papers

This title works! :)

Rank Is All You Need: Estimating the Trace of Powers of Density Matrices

Myeongjin Shin, Junseo Lee, Seungwoo Lee, Kabgyun Jeong

Aug 02 2024 quant-ph arXiv:2408.00314v1

Estimating the trace of powers of identical k density matrices (i.e., $\mathrm{Tr}(\rho^k)$) is a crucial subroutine for many applications such as calculating nonlinear functions of quantum states, preparing quantum Gibbs states, and mitigating quantum errors. Reducing the requisite number of qubits and gates is essential to fit a quantum algorithm onto near-term quantum devices. Inspired by the Newton-Girard method, we developed an algorithm that uses only $\mathcal{O}(r)$ qubits and $\mathcal{O}(r)$ multi-qubit gates, where r is the rank of ρ. We prove that the estimation of $\{\mathrm{Tr}(\rho^i)\}_{i=1}^r$ is sufficient for estimating the trace of powers with large $k > r$. With these advantages, our algorithm brings the estimation of the trace of powers closer to the capabilities of near-term quantum processors. We show that our results can be generalized for estimating $\mathrm{Tr}(M\rho^k)$, where M is an arbitrary observable, and demonstrate the advantages of our algorithm in several applications.

Overview

- **Trace of powers & Literature review**
- Mathematical intuitions
- Main results: algorithm, lemmas, theorems, corollaries
- Numerical simulations
- Applications
- Concluding remarks

Trace of powers

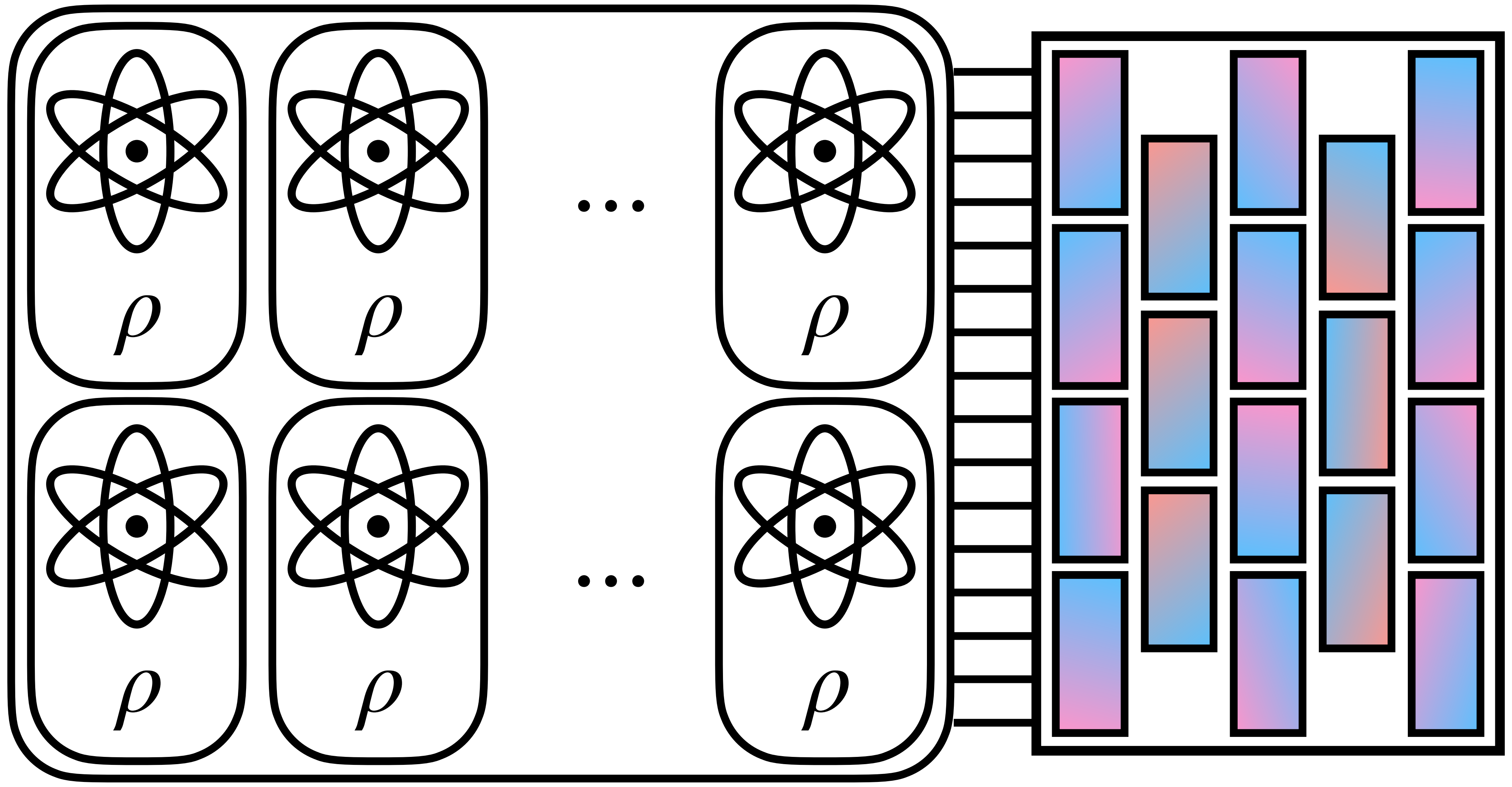
How can we estimate the value of $\mathrm{Tr}\left(\rho^k\right)$

when given access to copies of a quantum state ρ? (for large $k \in \mathbb{N}$)

Trace of powers

How can we estimate the value of $\mathrm{Tr}\left(\rho^k\right)$

when given access to copies of a quantum state ρ? (for large $k \in \mathbb{N}$)



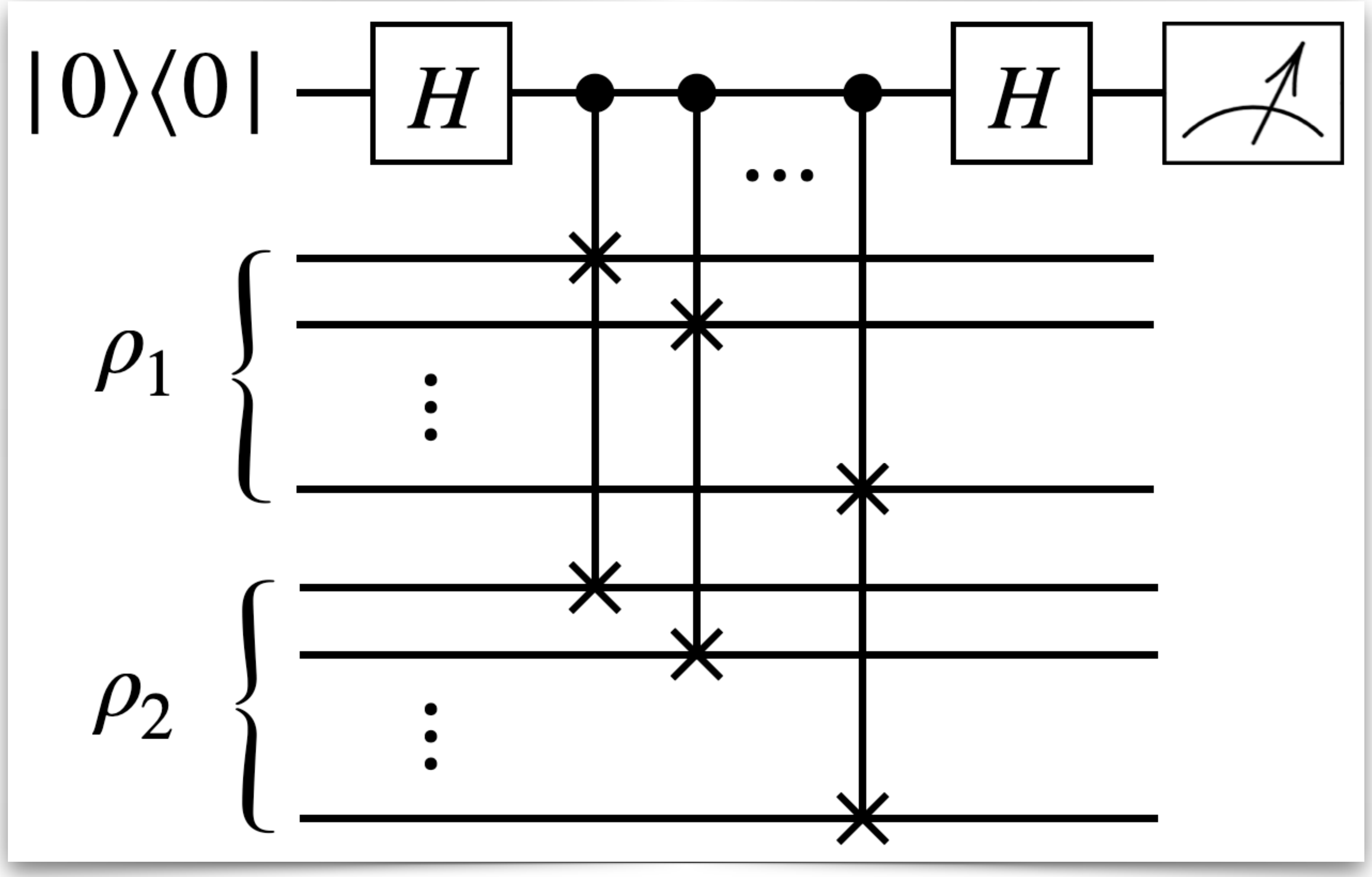
Given $\rho^{\otimes k}$

(efficient) quantum protocol
& some post processing

$\Rightarrow$ ε-approximate $\mathrm{Tr}\left(\rho^k\right)$

- Integer Rényi entropy
- Nonlinear function calculations
- Entanglement spectroscopy
- Quantum error mitigation
- Quantum Gibbs state preparation

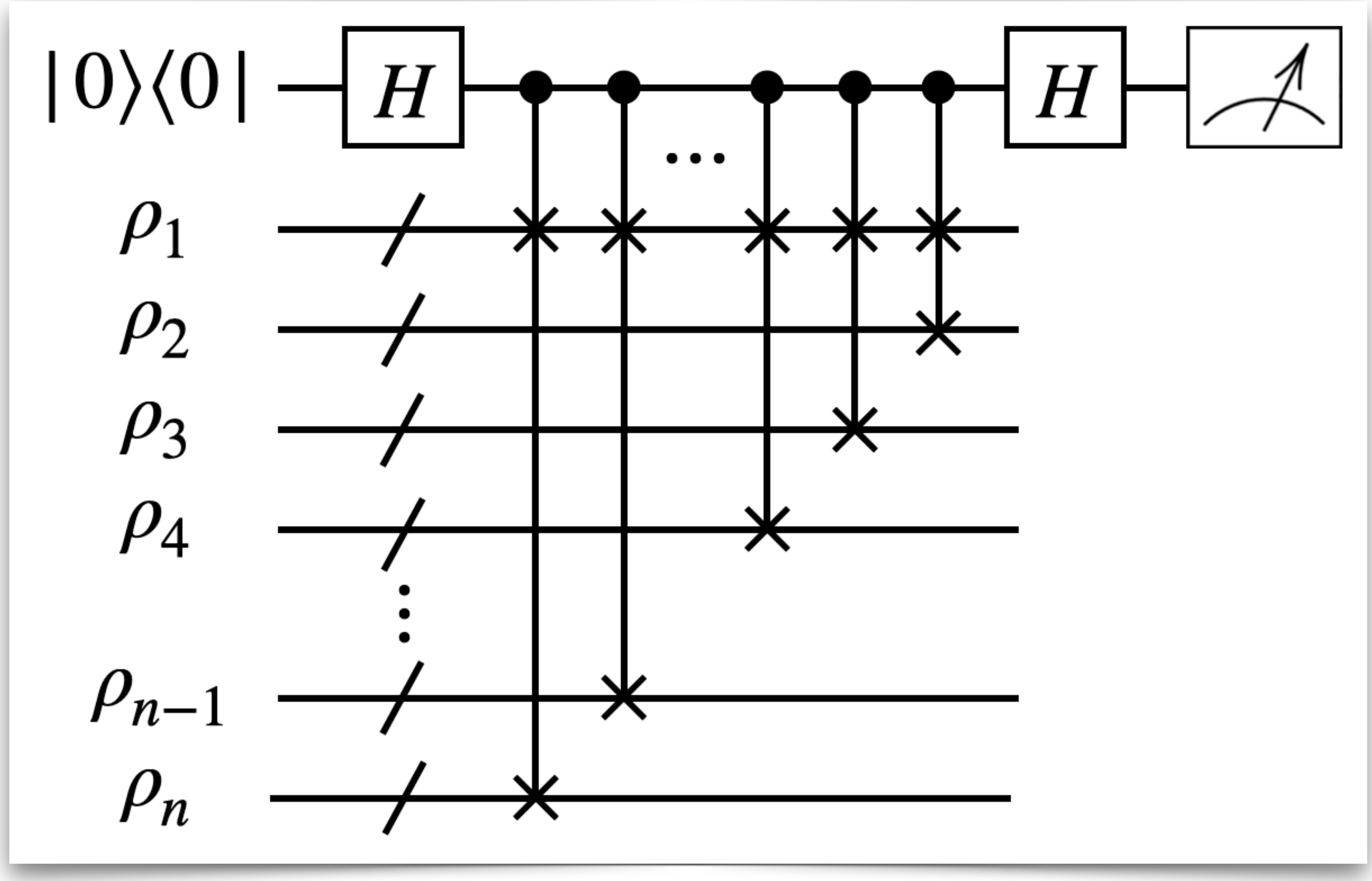
Swap test



$$\mathrm{Tr}\left(S\left(\rho_1 \otimes \rho_2\right)\right) = \mathrm{Tr}\left(\rho_1\rho_2\right)$$

S = swap operator

Generalized swap test



To estimate Tr $\left(\rho^k\right)$

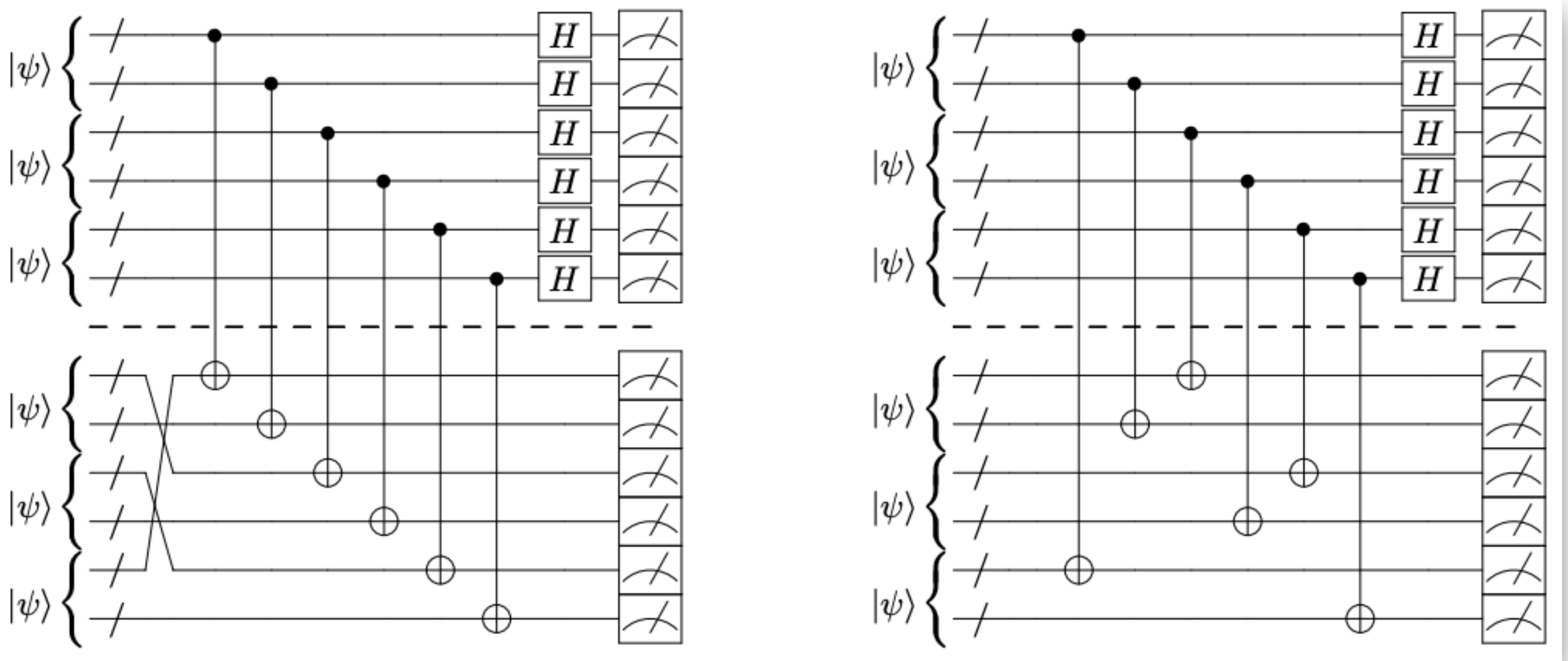
* Depth: $\mathcal{O}(k)$
* Width: $\mathcal{O}(k)$
* Copies: $\mathcal{O}(k)$
* Multi-qubit gates: $\mathcal{O}(k)$

$$\mathrm{Tr}\left(W^\pi\left(\rho_1\otimes\rho_2\otimes\cdots\otimes\rho_{n-1}\otimes\rho_n\right)\right)=\mathrm{Tr}\left(\rho_1\rho_2\cdots\rho_{n-1}\rho_n\right)$$

W^π = cyclic shift permutation operator

"Direct estimations of linear and nonlinear functionals of a quantum state". Physical Review Letters 88, 217901 (2002).

Two copy test



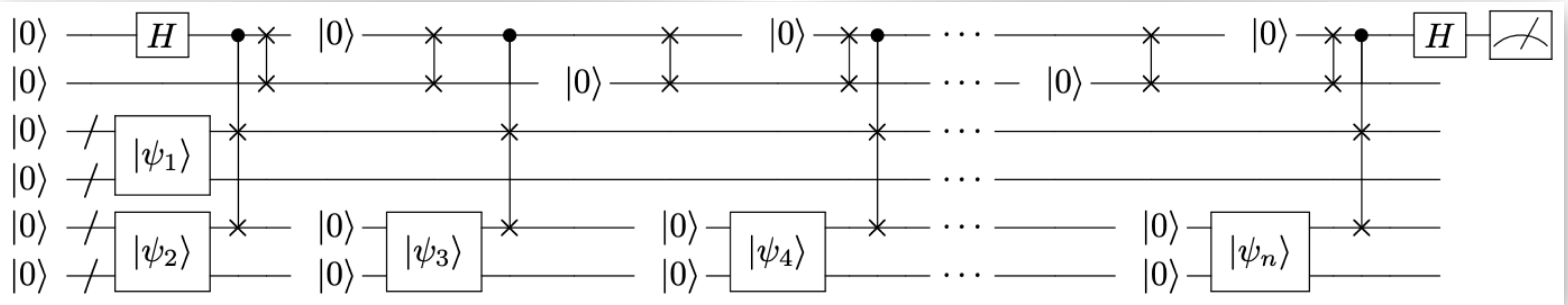
To estimate $\mathrm{Tr}\left(\rho^k\right)$

- **Depth: $\mathcal{O}(1)$**
- Width: $\mathcal{O}(k)$
- Copies: $\mathcal{O}(k)$
- Multi-qubit gates: $\mathcal{O}(k)$

* Note that original entangled pure state $|\psi_{AB}\rangle$ needed, where $\rho_A = \mathrm{Tr}_B\left(|\psi\rangle\langle\psi|_{AB}\right)$

"Entanglement spectroscopy with a depth-two quantum circuit". Journal of Physics A: Mathematical and Theoretical 52, 044001 (2019).

Qubit-efficient two copy test



To estimate $\mathrm{Tr}\left(\rho^k\right)$

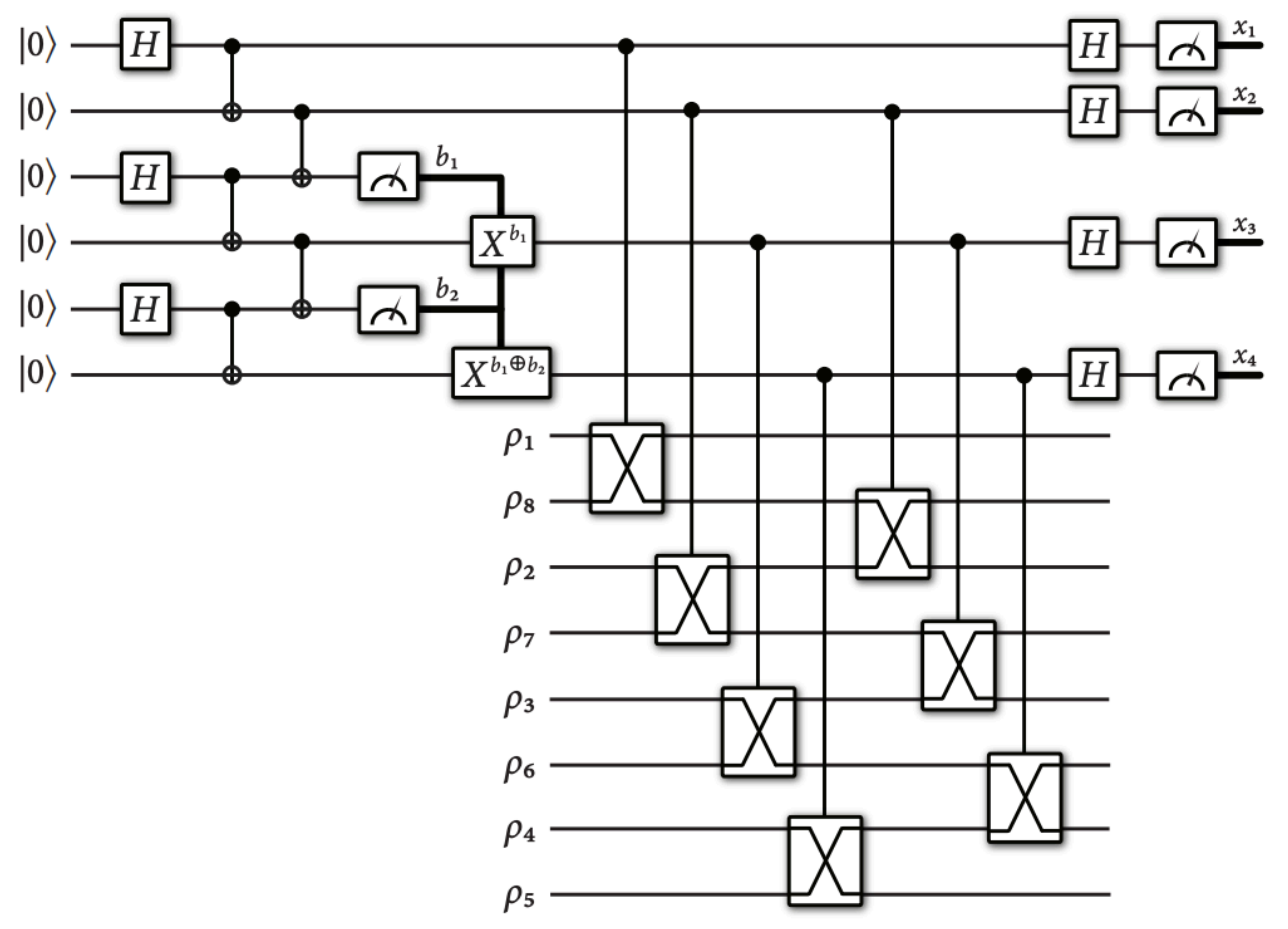
* Depth: $\mathcal{O}(k)$
* **Width: $\mathcal{O}(1)$**
* Copies: $\mathcal{O}(k)$
* Multi-qubit gates: $\mathcal{O}(k)$

* Using qubit-reset strategies

* Original entangled pure state $|\psi_{AB}\rangle$ needed

"Qubit- efficient entanglement spectroscopy using qubit resets". Quantum 5, 535 (2021).

Multivariate trace estimation algorithm



To estimate $\mathrm{Tr}\left(\rho^k\right)$

- **Depth: $\mathcal{O}(1)$**
- Width: $\mathcal{O}(k)$
- Copies: $\mathcal{O}(k)$
- Multi-qubit gates: $\mathcal{O}(k)$

* Inspired by Shor's error correction code

"Multivariate trace estimation in constant quantum depth". Quantum 8, 1220 (2024)

Comparison

Summary of resources required by different algorithms
to estimate the values of $\{\mathrm{Tr}(\rho^i)\}_{i=1}^k$ within an error margin of ϵ

Algorithm	# Depth	# Qubits	# CSWAP	# Copies	Original $\|\psi\rangle$
Generalized swap test	$\mathcal{O}(k)$	$\mathcal{O}(k)$	$\mathcal{O}(k)$	$\mathcal{O}\left(\frac{k^2}{\epsilon^2}\right)$	***NOT*** required
Hadamard test	$\mathcal{O}(k)$	$\mathcal{O}(k)$	$\mathcal{O}(k)$	$\mathcal{O}\left(\frac{k^2}{\epsilon^2}\right)$	Required
Two copy test	$\mathcal{O}(1)$	$\mathcal{O}(k)$	$\mathcal{O}(k)$	$\mathcal{O}\left(\frac{k^2}{\epsilon^2}\right)$	Required
Qubit-efficient two copy test	$\mathcal{O}(k)$	$\mathcal{O}(1)$	$\mathcal{O}(k)$	$\mathcal{O}\left(\frac{k^2}{\epsilon^2}\right)$	Required
Multivariate trace estimation	$\mathcal{O}(1)$	$\mathcal{O}(k)$	$\mathcal{O}(k)$	$\mathcal{O}\left(\frac{k^2}{\epsilon^2}\right)$	***NOT*** required
Ours (this work)	$\mathcal{O}(1)$	$\mathcal{O}(r)$	$\mathcal{O}(r)$	$\mathcal{O}\left(\frac{k^2 r^4 \ln^2 r}{\epsilon^2}\right)$, $p_1 \approx 1$ $\mathcal{O}\left(\frac{r^2 \ln^2 r}{\epsilon^2}\right)$, otherwise	***NOT*** required

Overview

- Trace of powers & Literature review
- **Mathematical intuitions**
- Main results: algorithm, lemmas, theorems, corollaries
- Numerical simulations
- Applications
- Concluding remarks

Intuition

Consider two quantum states: $\rho = \sum_{i=1}^{r} p_i |\psi_i\rangle\langle\psi_i|, \, \sigma = \sum_{i=1}^{r} q_i |\phi_i\rangle\langle\phi_i|$

(assume descending order $p_1 \geq p_2 \geq \ldots p_{r-1} \geq p_r \geq 0$ and also for q_i)

Intuition

Consider two quantum states: $\rho = \sum_{i=1}^{r} p_i |\psi_i\rangle\langle\psi_i|$, $\sigma = \sum_{i=1}^{r} q_i |\phi_i\rangle\langle\phi_i|$

(assume descending order $p_1 \geq p_2 \geq \ldots p_{r-1} \geq p_r \geq 0$ and also for q_i)

If $\begin{cases} \mathrm{Tr}(\rho^1) = \mathrm{Tr}(\sigma^1) \\ \mathrm{Tr}(\rho^2) = \mathrm{Tr}(\sigma^2) \\ \vdots \\ \mathrm{Tr}(\rho^{r-1}) = \mathrm{Tr}(\sigma^{r-1}) \\ \mathrm{Tr}(\rho^r) = \mathrm{Tr}(\sigma^r) \end{cases}$ then, $\begin{cases} p_1 = q_1 \\ p_2 = q_2 \\ \vdots \\ p_{r-1} = q_{r-1} \\ p_r = q_r \end{cases}$?

Intuition

Consider two quantum states: $\rho = \sum_{i=1}^{r} p_i |\psi_i\rangle\langle\psi_i|$, $\sigma = \sum_{i=1}^{r} q_i |\phi_i\rangle\langle\phi_i|$

(assume descending order $p_1 \geq p_2 \geq \ldots p_{r-1} \geq p_r \geq 0$ and also for q_i)

If $\begin{cases} \mathrm{Tr}(\rho^1) = \mathrm{Tr}(\sigma^1) \\ \mathrm{Tr}(\rho^2) = \mathrm{Tr}(\sigma^2) \\ \vdots \\ \mathrm{Tr}(\rho^{r-1}) = \mathrm{Tr}(\sigma^{r-1}) \\ \mathrm{Tr}(\rho^r) = \mathrm{Tr}(\sigma^r) \end{cases}$ then, $\begin{cases} p_1 = q_1 \\ p_2 = q_2 \\ \vdots \\ p_{r-1} = q_{r-1} \\ p_r = q_r \end{cases}$?

<u>**Answer: YES**</u>

Newton-Girard method

Key observation: exactly knowing $\{\mathrm{Tr}(\rho^i)\}_{i=1}^r$ is equivalent to knowing $\{p_i\}_{i=1}^r$.

Newton-Girard method

Key observation: exactly knowing $\{\mathrm{Tr}(\rho^i)\}_{i=1}^r$ is equivalent to knowing $\{p_i\}_{i=1}^r$.

Consider the equation having these eigenvalues as root in the form of

$$\prod_{m=1}^{r}(x - p_m) = 0.$$

Newton-Girard method

Key observation: exactly knowing $\{\mathrm{Tr}(\rho^i)\}_{i=1}^r$ is equivalent to knowing $\{p_i\}_{i=1}^r$.

Consider the equation having these eigenvalues as root in the form of

$$\prod_{m=1}^{r}(x - p_m) = 0.$$

The values of $\mathrm{Tr}(\rho^i)$ are now the i-th power sum of the roots. Denote the power sum as

$$P_i := \sum_{m=1}^{r} p_m^i = \mathrm{Tr}(\rho^i)\,.$$

Newton-Girard method

Simply expanding the terms to get: $\prod_{m=1}^{r}(x - p_m) = \sum_{k=0}^{r}(-1)^k a_k x^{r-k}$,

Newton-Girard method

Simply expanding the terms to get: $\prod_{m=1}^{r}(x - p_m) = \sum_{k=0}^{r}(-1)^k a_k x^{r-k}$,

where

$$
\begin{aligned}
a_0 &= 1,\\
a_1 &= p_1 + p_2 + \cdots + p_r = \sum_{1\le\alpha\le r} p_\alpha,\\
a_2 &= p_1p_2 + p_1p_3 + \cdots + p_{r-1}p_r = \sum_{1\le\alpha<\beta\le r} p_\alpha p_\beta,\\
a_3 &= \sum_{1\le\alpha<\beta<\gamma\le r} p_\alpha p_\beta p_\gamma,\\
&\vdots\\
a_r &= \prod_{i=1}^{r} p_i\,.
\end{aligned}
$$

Newton-Girard method

The Newton-Girard method states the relationship between

the elementary symmetric polynomials and the power sums recursively.

For all $r \geq k \geq 1, \;\; a_k = \dfrac{1}{k}\sum_{i=1}^{k}(-1)^{i-1}a_{k-i}P_i.$

Newton-Girard method

The Newton-Girard method states the relationship between

the elementary symmetric polynomials and the power sums recursively.

For all $r \geq k \geq 1, \quad a_k = \frac{1}{k} \sum_{i=1}^{k} (-1)^{i-1} a_{k-i} P_i.$

Uniquely defined $P_i = \mathrm{Tr}(\rho^i)$

Newton-Girard method

The Newton-Girard method states the relationship between

the elementary symmetric polynomials and the power sums recursively.

For all $r \geq k \geq 1, \quad a_k = \frac{1}{k} \sum_{i=1}^{k} (-1)^{i-1} a_{k-i} P_i.$

Also unique

Uniquely defined $P_i = \mathrm{Tr}(\rho^i)$

Newton-Girard method

The Newton-Girard method states the relationship between

the elementary symmetric polynomials and the power sums recursively.

For all $r \geq k \geq 1, \quad a_k = \frac{1}{k} \sum_{i=1}^{k} (-1)^{i-1} a_{k-i} P_i.$

Also unique

Uniquely defined $P_i = \mathrm{Tr}(\rho^i)$

Therefore, the set of eigenvalues is **uniquely determined**

as the roots of the equation $\prod_{m=1}^{r} (x - p_m)$.

Estimation with errors

Unfortunately in real-world situations,

we cannot exactly calculate the trace of powers.

Estimation with errors

Unfortunately in real-world situations,
we cannot exactly calculate the trace of powers.

Next challenge: If the error of estimated power sums $P_i = \mathrm{Tr}(\rho^i)$ is small, are the roots obtained by the Newton-Girard method close to the eigenvalues of ρ?

Estimation with errors

Unfortunately in real-world situations,
we cannot exactly calculate the trace of powers.

Next challenge: If the error of estimated power sums $P_i = \mathrm{Tr}(\rho^i)$ is small,
are the roots obtained by the Newton-Girard method close to the eigenvalues of ρ?

Answer: NO

Estimation with errors

Unfortunately in real-world situations,
we cannot exactly calculate the trace of powers.

Next challenge: If the error of estimated power sums $P_i = \mathrm{Tr}(\rho^i)$ is small,
are the roots obtained by the Newton-Girard method close to the eigenvalues of ρ?

Answer: NO

Counterexample – *Wilkinson's polynomial*

* The location of the roots can be very sensitive to perturbations
in the coefficients of the polynomial

Wilkinson's polynomial

$$w(x) = \prod_{i=1}^{20} (x - i) = (x - 1)(x - 2)\cdots(x - 20)$$

Expanding the polynomial, one finds: $w(x) = x^{20} - 210x^{19} + 20615x^{18} - \cdots$

Wilkinson's polynomial

$$w(x) = \prod_{i=1}^{20} (x - i) = (x - 1)(x - 2)\cdots(x - 20)$$

Expanding the polynomial, one finds: $w(x) = x^{20} - 210x^{19} + 20615x^{18} - \cdots$

If the coefficient of x^{19} is decreased from -210 by 2^{-23} to -210.0000001192

Wilkinson's polynomial

$$w(x) = \prod_{i=1}^{20} (x - i) = (x - 1)(x - 2)\cdots(x - 20)$$

Expanding the polynomial, one finds: $w(x) = x^{20} - 210x^{19} + 20615x^{18} - \cdots$

If the coefficient of x^{19} is decreased from -210 by 2^{-23} to -210.0000001192

The 20 roots become:

1.00000	2.00000	3.00000	4.00000	5.00000
6.00001	6.99970	8.00727	8.91725	20.84691
$10.09527 \pm 0.64350i$	$11.79363 \pm 1.65233i$	$13.99236 \pm 2.51883i$	$16.73074 \pm 2.81262i$	$19.50244 \pm 1.94033i$

Wilkinson's polynomial

$$w(x) = \prod_{i=1}^{20}(x - i) = (x - 1)(x - 2)\cdots(x - 20)$$

Expanding the polynomial, one finds: $w(x) = x^{20} - 210x^{19} + 20615x^{18} - \cdots$

If the coefficient of x^{19} is decreased from -210 by 2^{-23} to -210.0000001192

The 20 roots become:

1.00000	2.00000	3.00000	4.00000	5.00000
6.00001	6.99970	8.00727	8.91725	20.84691
$10.09527 \pm 0.64350i$	$11.79363 \pm 1.65233i$	$13.99236 \pm 2.51883i$	$16.73074 \pm 2.81262i$	$19.50244 \pm 1.94033i$

* Some of the roots are greatly displaced, even though the change to the coefficient is tiny and the original roots seem widely spaced.

Overview

- Trace of powers & Literature review
- Mathematical intuitions
- **Main results: algorithm, lemmas, theorems, corollaries**
- Numerical simulations
- Applications
- Concluding remarks

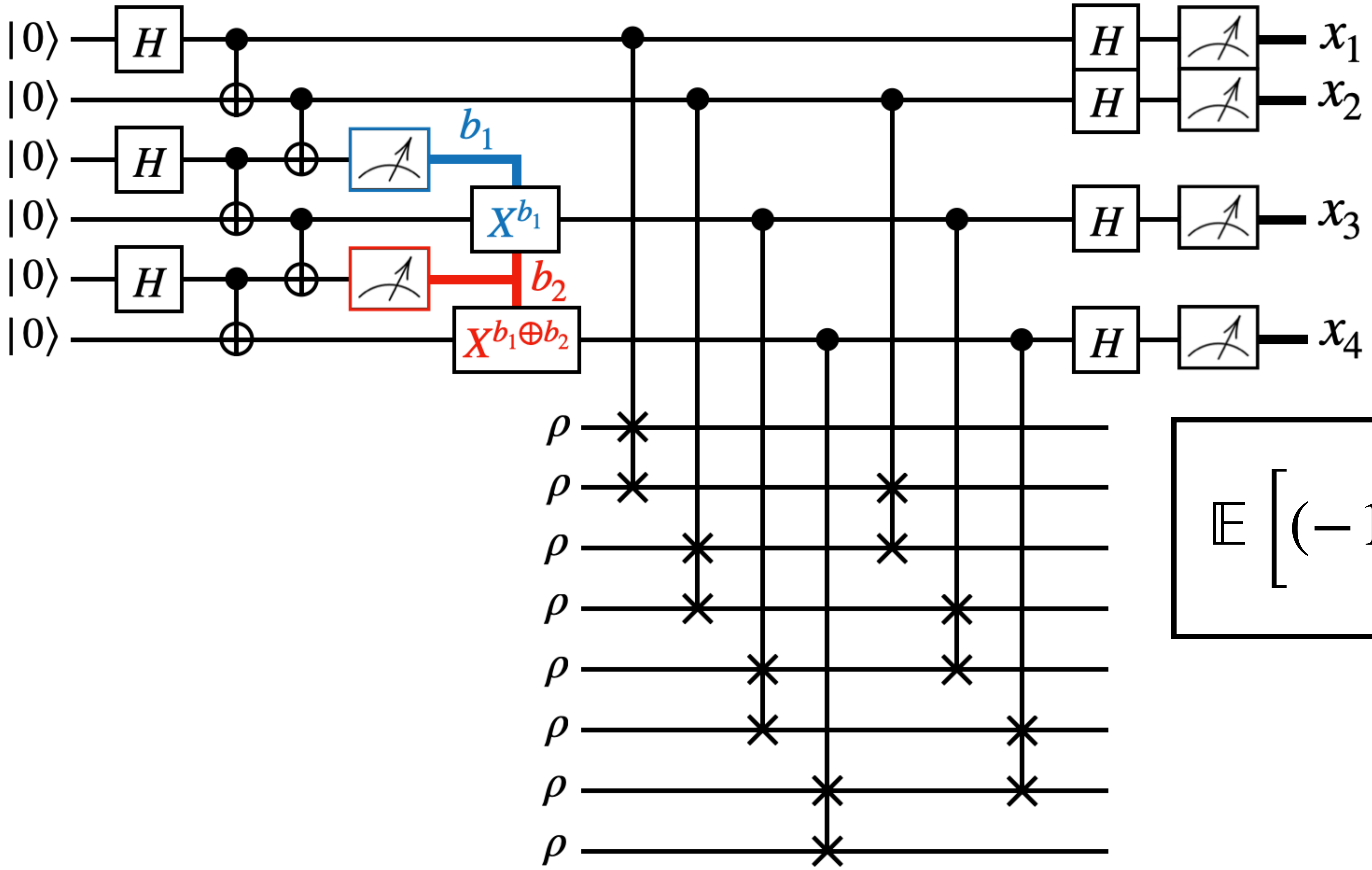
Main result: iterative $\mathrm{Tr}(\rho^i)$ estimation algorithm

[1] Estimate $P_i = \mathrm{Tr}(\rho^i)$ for $i = 1,2,\ldots,r$, using a constant-depth quantum circuit consisting of $\mathcal{O}(i)$ qubits and $\mathcal{O}(i)$ CSWAP operations using multivariate trace estimation, where r is the rank of ρ, and denote the estimated value as Q_i. $\rightarrow \; Q_1, \ldots, Q_r$

Main result: iterative $\mathrm{Tr}(\rho^i)$ estimation algorithm

[1] Estimate $P_i = \mathrm{Tr}(\rho^i)$ for $i = 1,2,\ldots,r$, using a constant-depth quantum circuit consisting of $\mathcal{O}(i)$ qubits and $\mathcal{O}(i)$ CSWAP operations using multivariate trace estimation, where r is the rank of ρ, and denote the estimated value as Q_i. $\rightarrow\ Q_1, \ldots, Q_r$

Example ($i = 8$)



$$\mathbb{E}\left[(-1)^{\sum_{i=1}^{4} x_i}\right] = \mathrm{Re}\left(\mathrm{Tr}(\rho^8)\right)$$

Main result: iterative $\mathrm{Tr}(\rho^i)$ estimation algorithm

[1] Estimate $P_i = \mathrm{Tr}(\rho^i)$ for $i = 1,2,\ldots,r$, using a constant-depth quantum circuit consisting of $\mathcal{O}(i)$ qubits and $\mathcal{O}(i)$ CSWAP operations using multivariate trace estimation, where r is the rank of ρ, and denote the estimated value as Q_i.

$\rightarrow\ Q_1, \ldots, Q_r$

[2] Calculate the elementary symmetric polynomial $b_i = \frac{1}{i}\sum_{\ell=1}^{i}(-1)^{\ell-1} b_{i-\ell} Q_\ell,\ b_1 = 1,\ 1 \le i \le r$.

$\rightarrow\ b_1, \ldots, b_r$

Main result: iterative $\mathrm{Tr}(\rho^i)$ estimation algorithm

[1] Estimate $P_i = \mathrm{Tr}(\rho^i)$ for $i = 1,2,\ldots,r$, using a constant-depth quantum circuit consisting of $\mathcal{O}(i)$ qubits and $\mathcal{O}(i)$ CSWAP operations using multivariate trace estimation, where r is the rank of ρ, and denote the estimated value as Q_i.

$\rightarrow\ Q_1, \ldots, Q_r$

[2] Calculate the elementary symmetric polynomial $b_i = \frac{1}{i}\sum_{\ell=1}^{i}(-1)^{\ell-1}b_{i-\ell}Q_\ell,\ b_1 = 1,\ 1 \le i \le r$.

$\rightarrow\ b_1, \ldots, b_r$

[3] Calculate the estimated value Q_i $(i > r)$ by $Q_i = \sum_{\ell=1}^{r}(-1)^{\ell-1}b_\ell Q_{i-\ell} \sim \mathrm{Tr}(\rho^i)$.

$\rightarrow\ Q_{r+1}, \ldots$

Main result: iterative $\mathrm{Tr}(\rho^i)$ estimation algorithm

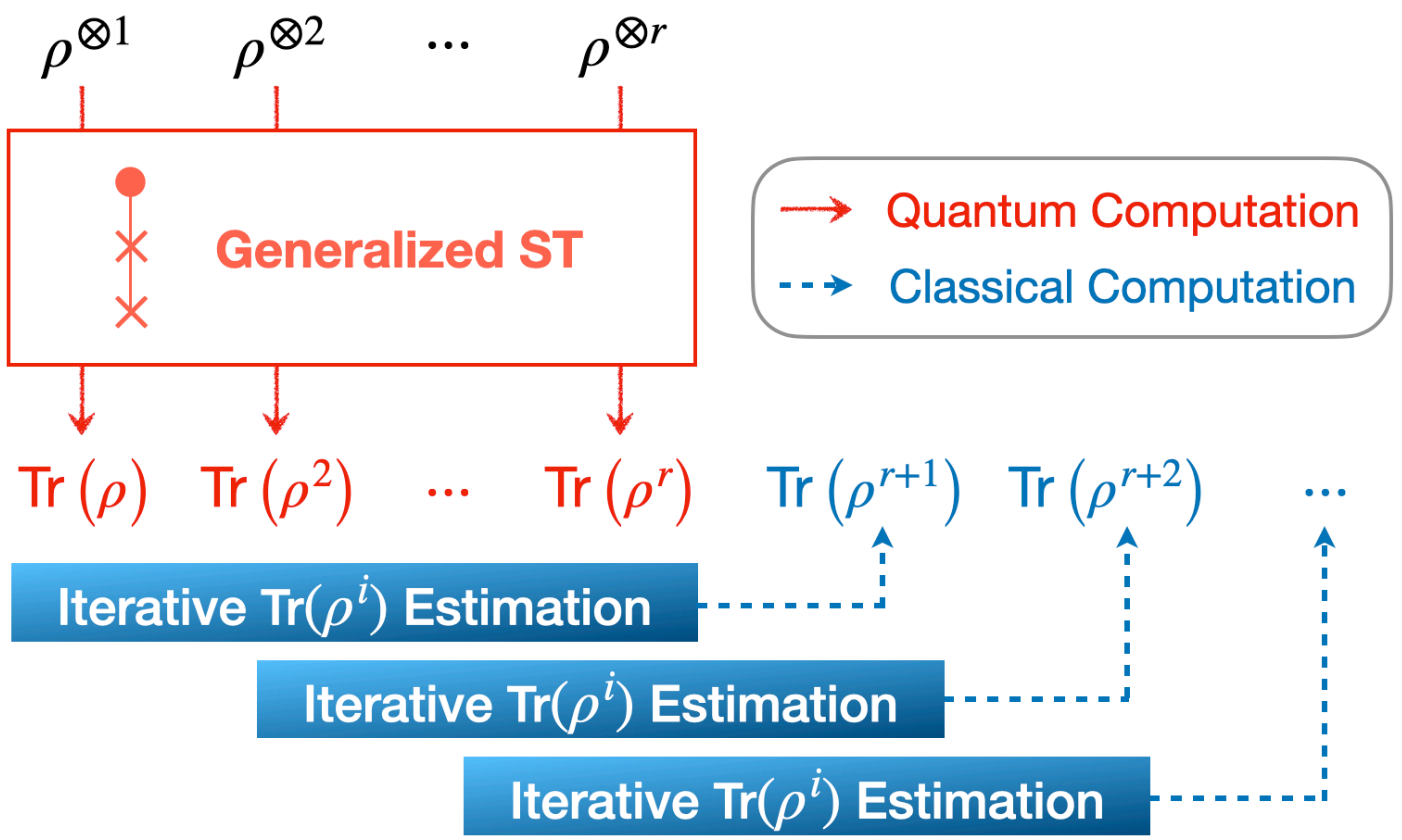
[1] Estimate $P_i = \mathrm{Tr}(\rho^i)$ for $i = 1,2,\ldots,r$, using a constant-depth quantum circuit consisting of $\mathcal{O}(i)$ qubits and $\mathcal{O}(i)$ CSWAP operations using multivariate trace estimation, where r is the rank of ρ, and denote the estimated value as Q_i. $\rightarrow\ Q_1,\ldots,Q_r$

[2] Calculate the elementary symmetric polynomial $b_i = \frac{1}{i}\sum_{\ell=1}^{i}(-1)^{\ell-1}b_{i-\ell}Q_\ell,\ b_1 = 1,\ 1 \le i \le r$. $\rightarrow\ b_1,\ldots,b_r$

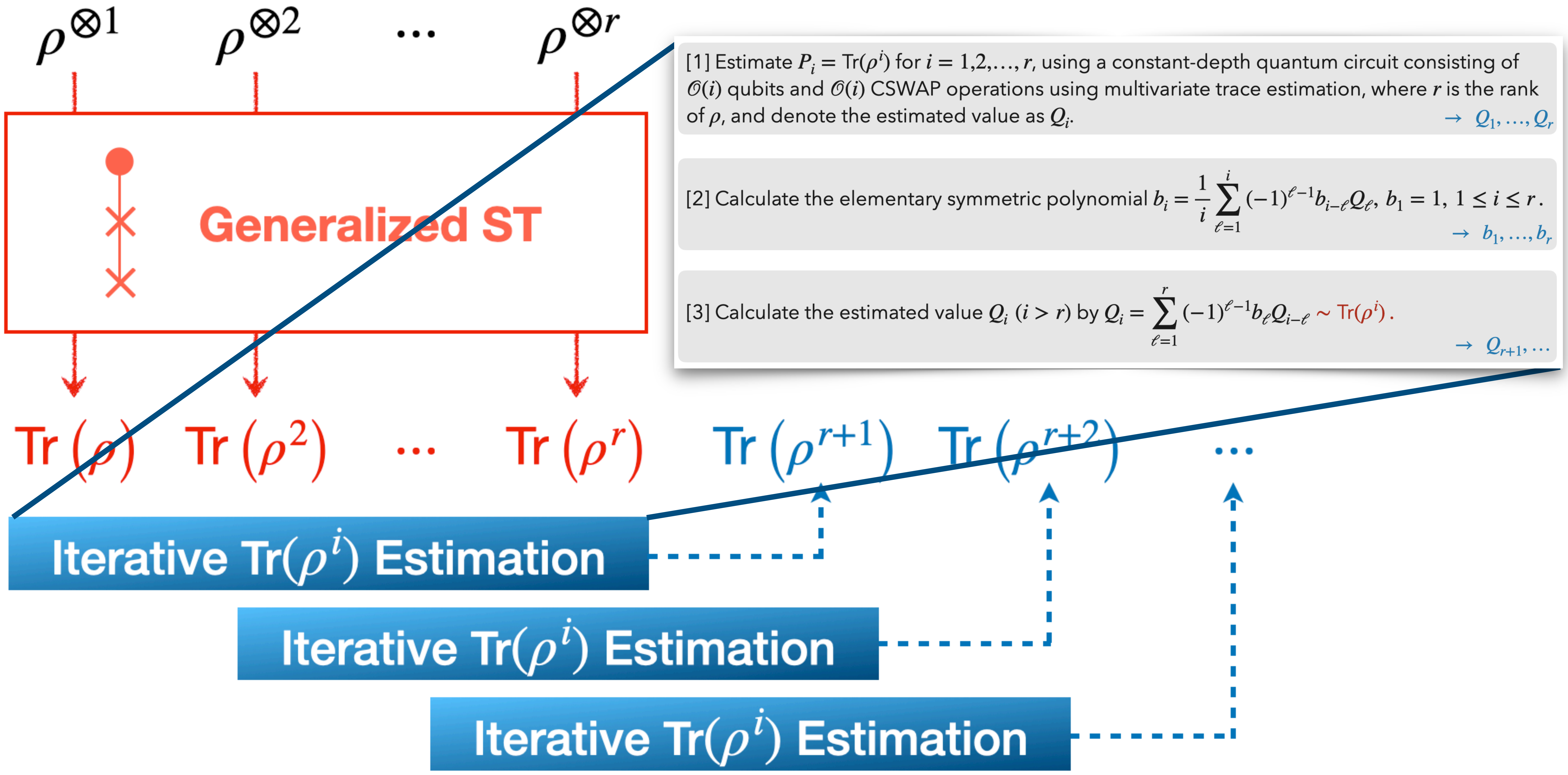
[3] Calculate the estimated value Q_i $(i > r)$ by $Q_i = \sum_{\ell=1}^{r}(-1)^{\ell-1}b_\ell Q_{i-\ell} \sim \mathrm{Tr}(\rho^i)$. $\rightarrow\ Q_{r+1},\ldots$

The output of our algorithm Q_i guarantees an *ϵ-approximate trace of powers* $\mathrm{Tr}(\rho^i)$

Main result: iterative $\mathrm{Tr}(\rho^i)$ estimation algorithm



Main result: iterative $\mathrm{Tr}(\rho^i)$ estimation algorithm



Rank is all you need — Lemma 1

- $a_k, b_k =$ the elementary symmetric polynomials corresponding to each P_i and Q_i.
- Q_i is defined as the estimated value of $P_i = \mathrm{Tr}(\rho^i)$ on a quantum device for $i \leq r$ otherwise $Q_i = \sum_{\ell=1}^{r} (-1)^{\ell-1} b_\ell Q_{i-\ell}$.

Rank is all you need – Lemma 1

- $a_k, b_k =$ the elementary symmetric polynomials corresponding to each P_i and Q_i.
- Q_i is defined as the estimated value of $P_i = \mathrm{Tr}(\rho^i)$ on a quantum device for $i \le r$ otherwise $Q_i = \sum_{\ell=1}^{r} (-1)^{\ell-1} b_\ell Q_{i-\ell}$.

Lemma 1

Let $t_k := b_k - a_k$, then the following holds: $|t_k| \le \sum_{j=1}^{k} \frac{|\epsilon_j|}{j}$

where $\epsilon_j = Q_j - P_j$ is the error that occurred by the estimation of $P_j = \mathrm{Tr}(\rho^j)$.

Rank is all you need – Lemma 1

- $a_k, b_k =$ the elementary symmetric polynomials corresponding to each P_i and Q_i.
- Q_i is defined as the estimated value of $P_i = \mathrm{Tr}(\rho^i)$ on a quantum device for $i \leq r$ otherwise $Q_i = \sum_{\ell=1}^{r} (-1)^{\ell-1} b_\ell Q_{i-\ell}$.

Lemma 1

Let $t_k := b_k - a_k$, then the following holds: $|t_k| \leq \sum_{j=1}^{k} \frac{|\epsilon_j|}{j}$

where $\epsilon_j = Q_j - P_j$ is the error that occurred by the estimation of $P_j = \mathrm{Tr}(\rho^j)$.

Proof. By strong mathematical induction logic.

Rank is all you need – Theorem 1

Theorem 1

Suppose that

$$\varepsilon_i := |\epsilon_i| = |P_i - Q_i| < \frac{\epsilon}{2T \ln r}$$

holds for $i = 1,2,\ldots,r$, where T is defined as:

$$T = \sum_{i=1}^{r} \frac{p_i(1 - p_i^k)(1 - p_i^r)}{(1 - p_i)^2} \leq kr\,.$$

Then the following relation always holds:

$$|\epsilon_i| = |P_i - Q_i| \leq \epsilon$$

for $i = 1,2,\ldots,k$.

Rank is all you need – Theorem 1

Theorem 1

Suppose that

$$\varepsilon_i := |\epsilon_i| = |P_i - Q_i| < \frac{\epsilon}{2T \ln r}$$

holds for $i = 1,2,\ldots,r$, where T is defined as:

$$T = \sum_{i=1}^{r} \frac{p_i(1 - p_i^k)(1 - p_i^r)}{(1 - p_i)^2} \leq kr\,.$$

Then the following relation always holds:

$$|\epsilon_i| = |P_i - Q_i| \leq \epsilon$$

for $i = 1,2,\ldots,k$.

Proof. By applying **Lemma 1** + long calculation with some mathematical trick.

Rank is all you need – Corollary 1

Corollary 1

To estimate Tr(ρ^i) for all $i \leq k$ within an additive error of ϵ and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$, it is necessary to estimate each Tr(ρ^j) for $j \leq r$ within an additive error of ε_j, as defined in **Theorem 1.** This can be achieved by using

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j)$ qubits and $\mathcal{O}(j)$ CSWAP operations. Here, T is defined in **Theorem 1.**

Rank is all you need – Corollary 1

Corollary 1

To estimate $\mathrm{Tr}(\rho^i)$ for all $i \leq k$ within an additive error of ϵ and with a success probability of at least $1 - \delta$, where $\delta \in (0,1)$, it is necessary to estimate each $\mathrm{Tr}(\rho^j)$ for $j \leq r$ within an additive error of ε_j, as defined in **Theorem 1.** This can be achieved by using

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2} \ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j)$ qubits and $\mathcal{O}(j)$ CSWAP operations. Here, T is defined in **Theorem 1.**

* Note that T scales from a constant to kr, mainly depending on the largest eigenvalue p_1 of ρ.
* If p_1 is not close to 1 (e.g., $p_1 = 0.5$), then $T = \mathcal{O}(1)$. This implies that if ρ is far from a pure state, then T is a constant value.

Rank is all you need – Theorem 2

Extension: Estimating $\mathrm{Tr}(M\rho^k)$, the trace of powers with arbitrary observables M.

Rank is all you need – Theorem 2

Extension: Estimating $\mathrm{Tr}(M\rho^k)$, the trace of powers with arbitrary observables M.

Theorem 2

Suppose that $\varepsilon_{i,M} := |\epsilon_{i,M}| = |P_{i,M} - Q_{i,M}| < \frac{\epsilon}{2}$, and $\varepsilon_i = |\epsilon_i| = |P_i - Q_i| < \frac{\epsilon}{2T\|M\|_\infty \ln r}$, holds for $i = 1,2,\ldots,r$, where the operator norm $\|M\|_\infty$ is defined corresponding to the ∞-norm for vectors $\|x\|_\infty$, as $\|M\|_\infty = \sup_{x\neq 0} \frac{\|Mx\|_\infty}{\|x\|_\infty}$, $T = \sum_{i=1}^{r} \frac{p_i(1-p_i^k)(1-p_i^r)}{(1-p_i)^2} \leq kr$.

Then the following holds:

$$|\epsilon_{i,M}| = |P_{i,M} - Q_{i,M}| \leq \epsilon$$

for $i = 1,2,\ldots,k$.

Rank is all you need – Corollary 2

Corollary 2

Suppose there is an efficient decomposition $M = \sum_{\ell=1}^{N_M} a_\ell P_\ell$, where $a_\ell \in \mathbb{R}$ and $P_\ell = \sigma_{\ell_1} \otimes \cdots \otimes \sigma_{\ell_n}$ are tensor products of Pauli operators $\sigma_{\ell_1}, \ldots, \sigma_{\ell_n} \in \{\sigma_x, \sigma_y, \sigma_z, I\}$. The quantity $\sum_{\ell=1}^{N_M} |a_\ell| = \mathcal{O}(c)$ is bounded by a constant c.

To estimate $\mathrm{Tr}(M\rho^i)$ for all $i \leq k$ within an additive error of ϵ and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$, it is necessary to estimate each $\mathrm{Tr}(M\rho^j)$ for $j \leq r$ within an additive error of $\varepsilon_{j,M}$.

This can be achieved by using $\mathcal{O}\left(\frac{c^2 N_M}{\epsilon^2} \ln\left(\frac{1}{\delta}\right)\right)$ runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j)$ qubits and $\mathcal{O}(j)$ CSWAP operations, and estimating each $\mathrm{Tr}(\rho^{j'})$ for $j' \leq r$ within an additive error of $\varepsilon_{j'}$, by using $\mathcal{O}\left(\frac{T^2}{\epsilon^2} \ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$ runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j')$ qubits and $\mathcal{O}(j')$ CSWAP operations. Here, $\varepsilon_{j,M}, \varepsilon_{j'}$ and T are defined in **Theorem 2.**

(Effective) Rank is all you need – Lemma 2

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

* If p_1 is not close to 1 (e.g., $p_1 = 0.5$), then $T = \mathcal{O}(1)$. This implies that if ρ is far from a pure state, then T is a constant value.
* ***Our algorithm may perform suboptimally on pure states.***

(Effective) Rank is all you need – Lemma 2

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

* If p_1 is not close to 1 (e.g., $p_1 = 0.5$), then $T = \mathcal{O}(1)$. This implies that if ρ is far from a pure state, then T is a constant value.
* ***Our algorithm may perform suboptimally on pure states.***

Definition (Effective rank)

Effective rank with error ϵ is defined as the minimum value r_ϵ, which satisfies $\sum_{i=1}^{r_\epsilon} p_i > 1 - \epsilon$.

(Effective) Rank is all you need – Lemma 2

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

* If p_1 is not close to 1 (e.g., $p_1 = 0.5$), then $T = \mathcal{O}(1)$. This implies that if ρ is far from a pure state, then T is a constant value.
* ***Our algorithm may perform suboptimally on pure states.***

Definition (Effective rank)

Effective rank with error ϵ is defined as the minimum value r_ϵ, which satisfies $\sum_{i=1}^{r_\epsilon} p_i > 1 - \epsilon$.

Lemma 2

For every integer $k \geq 2$, the following holds: $\left|\sum_{i=1}^{r} p_i^k - \sum_{i=1}^{r_\epsilon} p_i^k\right| < \epsilon^2$.

(Effective) Rank is all you need – Lemma 2

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

* If p_1 is not close to 1 (e.g., $p_1 = 0.5$), then $T = \mathcal{O}(1)$. This implies that if ρ is far from a pure state, then T is a constant value.
* ***Our algorithm may perform suboptimally on pure states.***

Definition (Effective rank)

Effective rank with error ϵ is defined as the minimum value r_ϵ, which satisfies $\sum_{i=1}^{r_\epsilon} p_i > 1 - \epsilon$.

Lemma 2

For every integer $k \geq 2$, the following holds: $\left|\sum_{i=1}^{r} p_i^k - \sum_{i=1}^{r_\epsilon} p_i^k\right| < \epsilon^2$.

Lemma 2 suggests that we only need $\{\mathrm{Tr}(\rho^i)\}_{i=1}^{r_\epsilon}$ for the estimation.

(Effective) Rank is all you need – Lemma 2

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

* If p_1 is not close to 1 (e.g., $p_1 = 0.5$), then $T = \mathcal{O}(1)$. This implies that if ρ is far from a pure state, then T is a constant value.
* ***Our algorithm may perform suboptimally on pure states.***

Definition (Effective rank)

Effective rank with error ϵ is defined as the minimum value r_ϵ, which satisfies $\sum_{i=1}^{r_\epsilon} p_i > 1 - \epsilon$.

Lemma 2

For every integer $k \geq 2$, the following holds: $\left|\sum_{i=1}^{r} p_i^k - \sum_{i=1}^{r_\epsilon} p_i^k\right| < \epsilon^2$.

Lemma 2 suggests that we only need $\{\mathrm{Tr}(\rho^i)\}_{i=1}^{r_\epsilon}$ for the estimation.

As the maximum eigenvalue of ρ approaches 1, the r_ϵ decreases!

(Possible to resolve issues with pure states)

Overview

- Trace of powers & Literature review
- Mathematical intuitions
- Main results: algorithm, lemmas, theorems, corollaries
- **Numerical simulations**
- Applications
- Concluding remarks

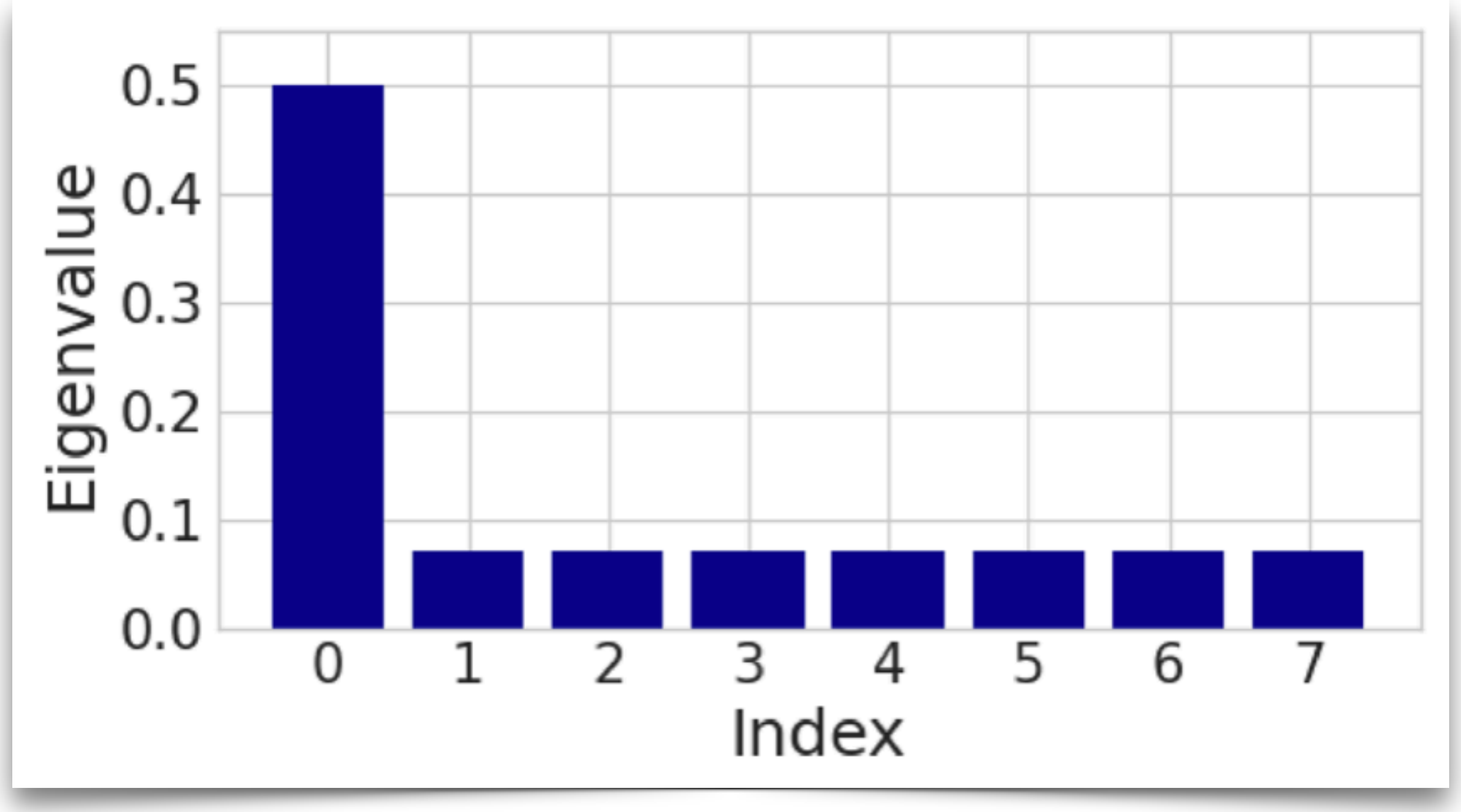
Reduction of required runs

Analyze the behavior of $T = \sum_{i=1}^{r} \frac{p_i(1-p_i^k)(1-p_i^r)}{(1-p_i)^2}$ as the eigenvalues of ρ change.

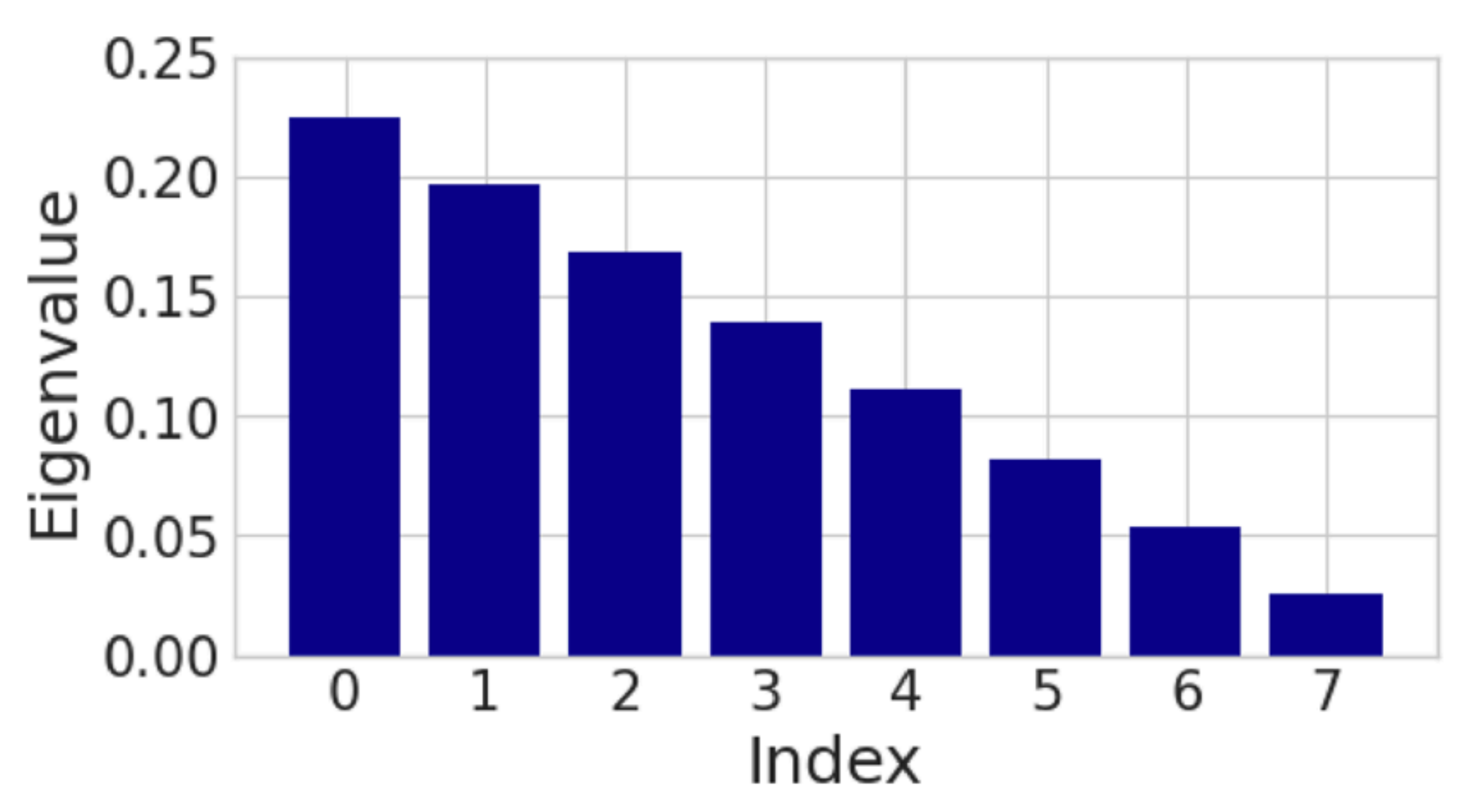
Reduction of required runs

Analyze the behavior of $T = \sum_{i=1}^{r} \frac{p_i(1-p_i^k)(1-p_i^r)}{(1-p_i)^2}$ as the eigenvalues of ρ change.

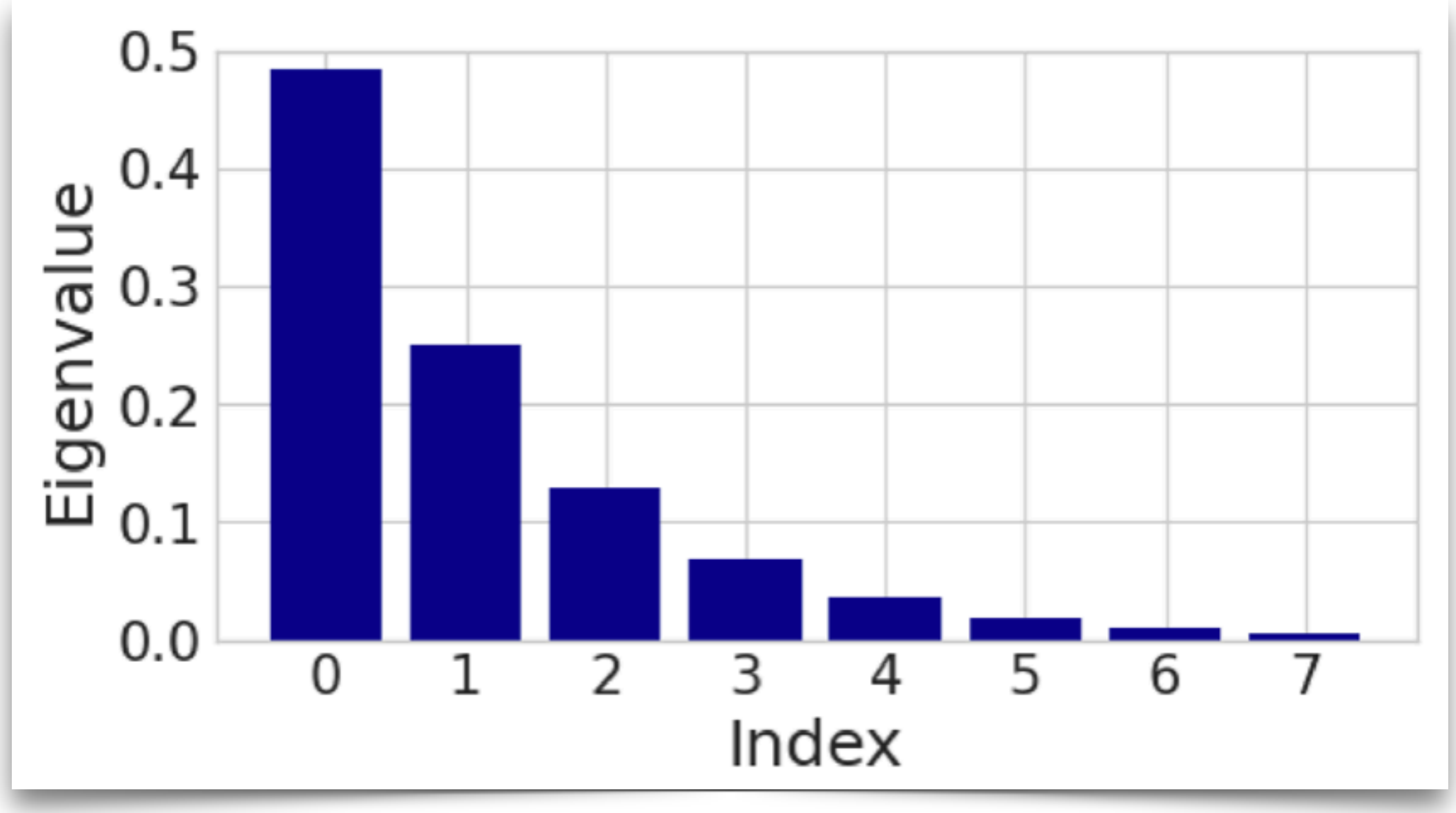
Since we don't know the exact distribution of the eigenvalues of an arbitrary density matrix ρ, we consider several typical cases.



Uniform non-maximum

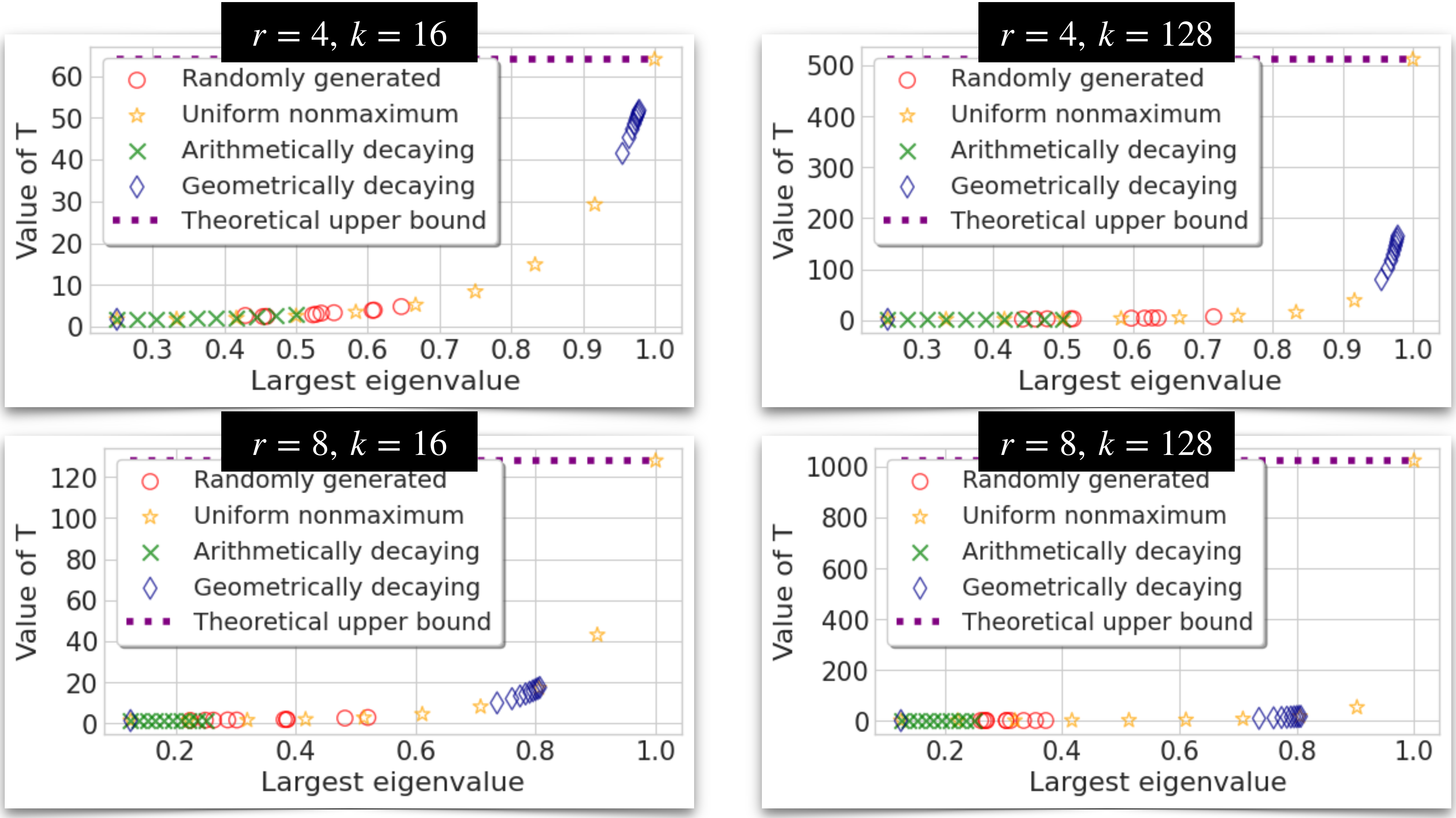


Arithmetically decaying

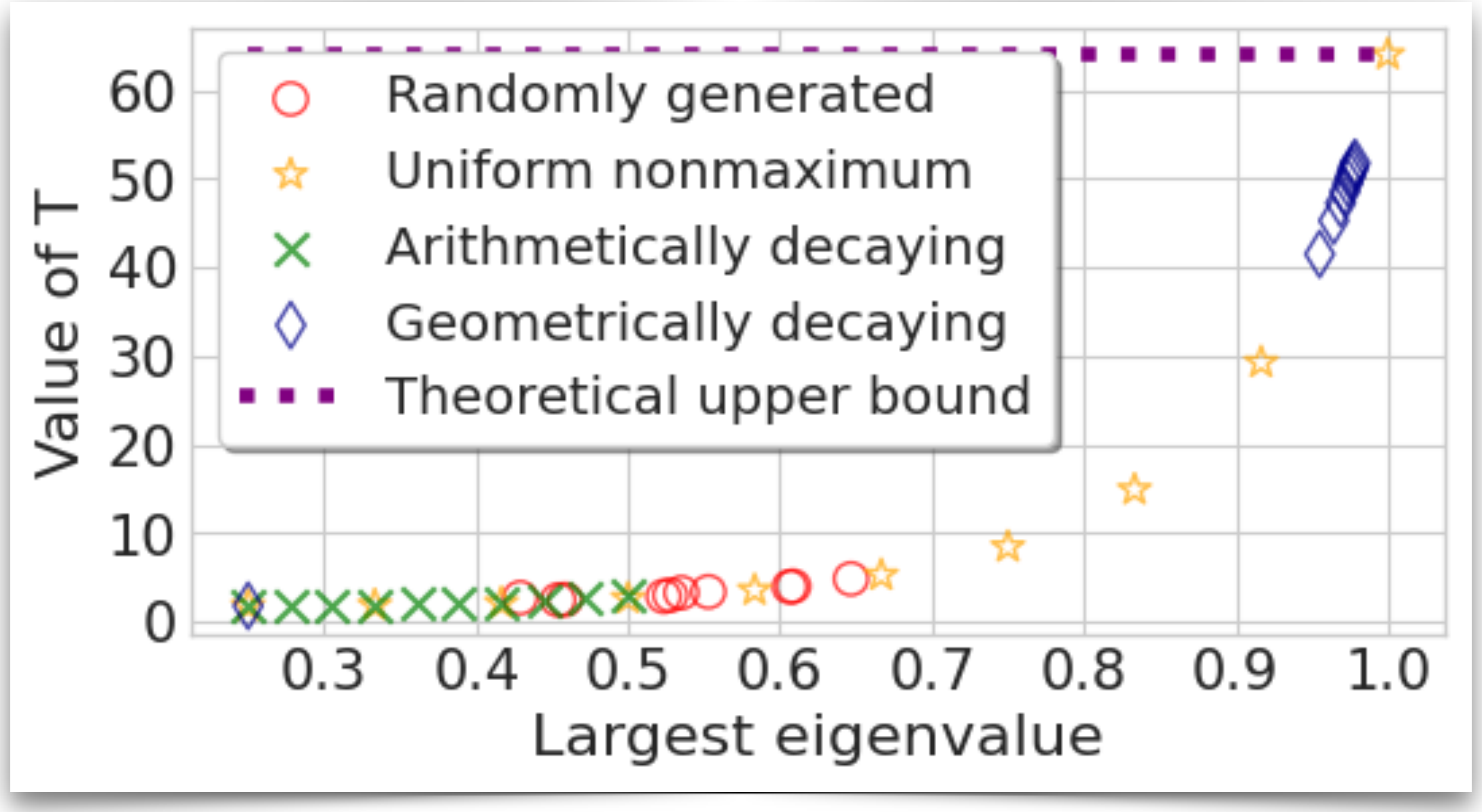


Geometrically decaying

Reduction of required runs

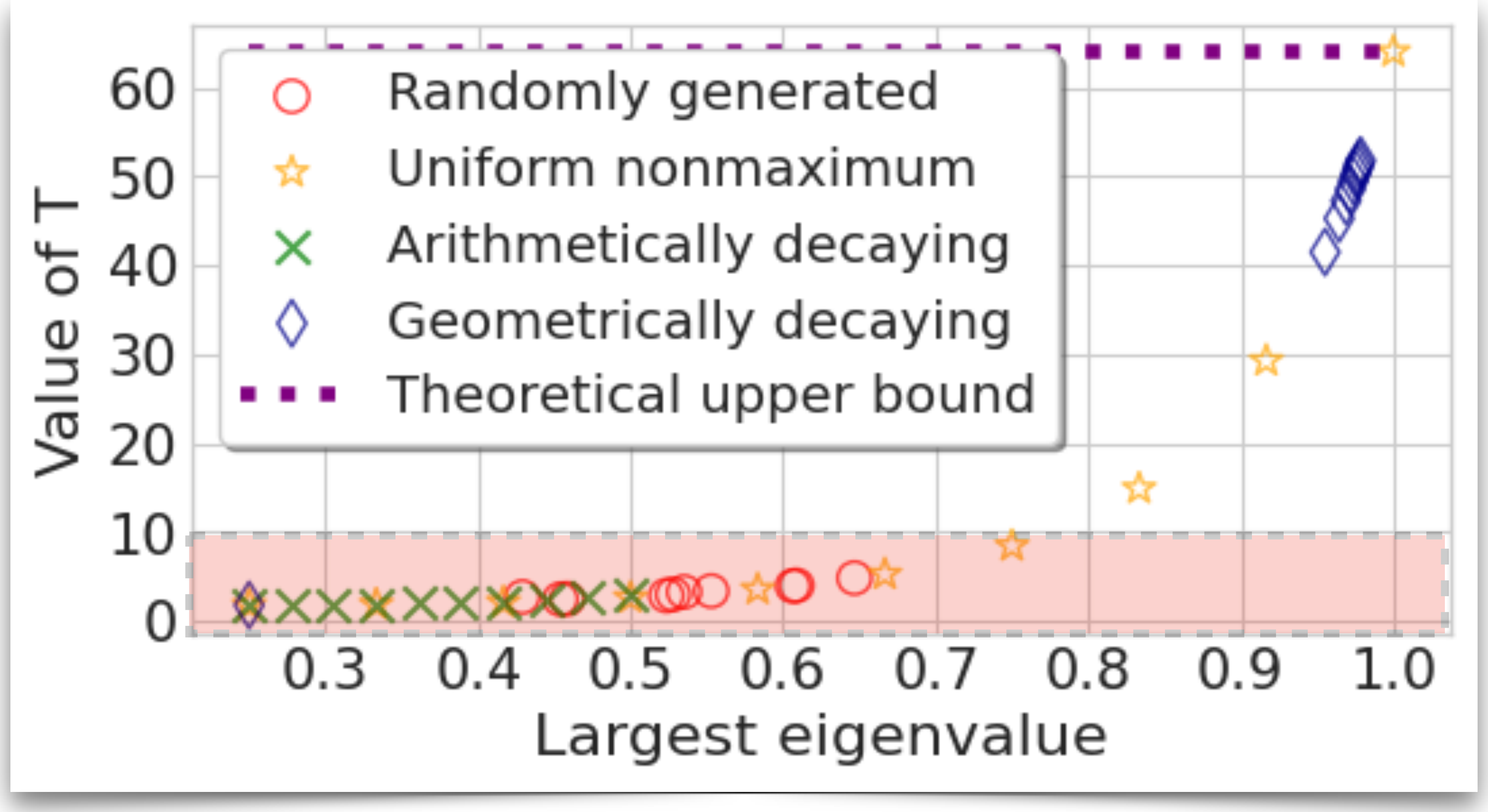


Reduction of required runs



$$T = \sum_{i=1}^{r} \frac{p_i(1-p_i^k)(1-p_i^r)}{(1-p_i)^2}$$

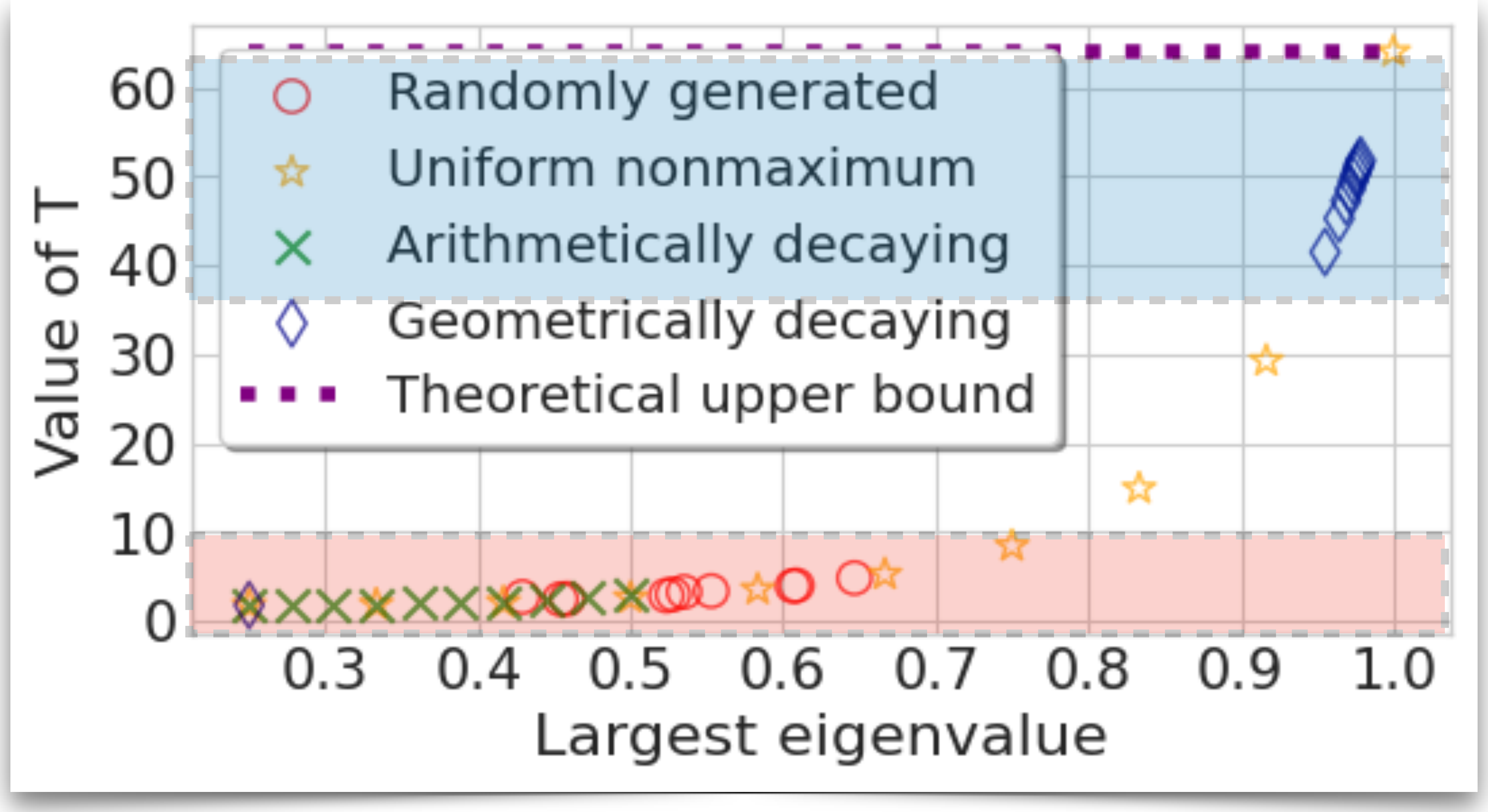
Reduction of required runs



~ constant

$$\mathcal{O}(1) \leq T = \sum_{i=1}^{r} \frac{p_i(1 - p_i^k)(1 - p_i^r)}{(1 - p_i)^2}$$

Reduction of required runs

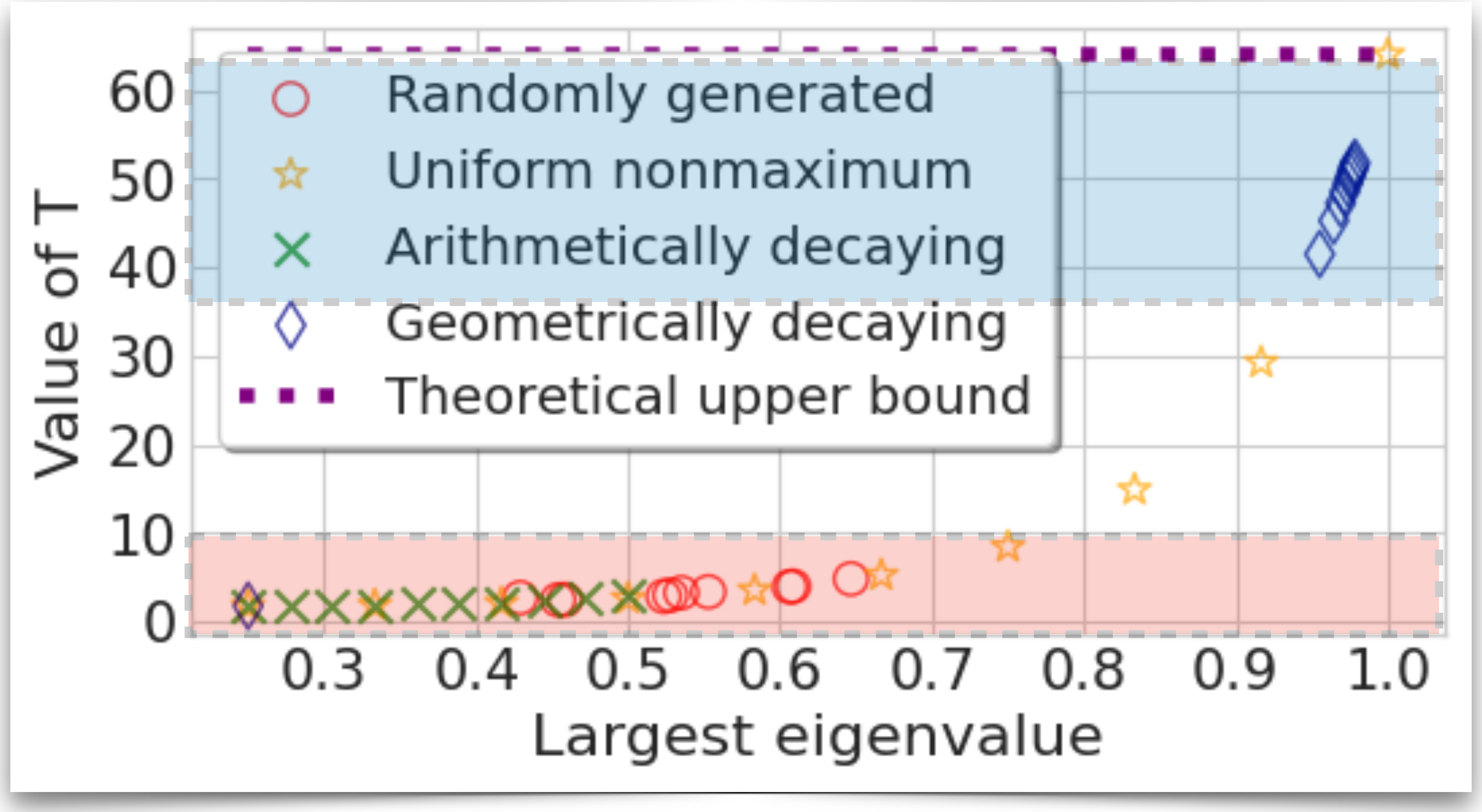


$\sim \mathcal{O}(kr)$

$$\mathcal{O}(1) \leq T = \sum_{i=1}^{r} \frac{p_i(1-p_i^k)(1-p_i^r)}{(1-p_i)^2} \leq kr$$

$\sim$ constant

Reduction of required runs



$\sim \mathcal{O}(kr)$

$\sim$ constant

$$\mathcal{O}(1) \leq T = \sum_{i=1}^{r} \frac{p_i(1-p_i^k)(1-p_i^r)}{(1-p_i)^2} \leq kr$$

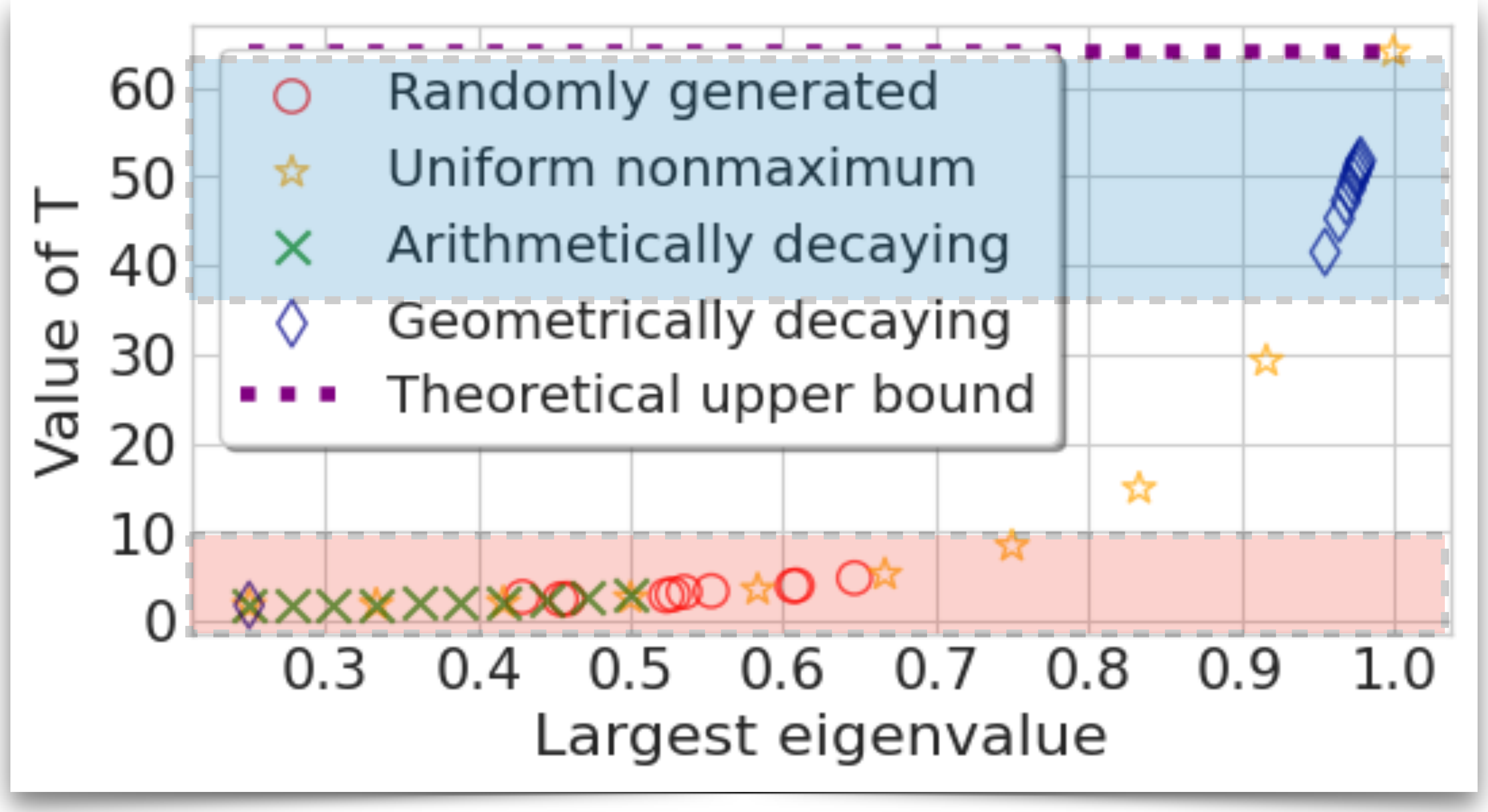
Corollary 1

To estimate $\mathrm{Tr}(\rho^i)$ for all $i \leq k$ within an additive error of ϵ and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$, it is necessary to estimate each $\mathrm{Tr}(\rho^j)$ for $j \leq r$ within an additive error of ϵ_j, as defined in **Theorem 1.** This can be achieved by using

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j)$ qubits and $\mathcal{O}(j)$ CSWAP operations. Here, T is defined in **Theorem 1.**

Reduction of required runs



$\sim \mathcal{O}(kr)$

$\sim$ constant

$$\mathcal{O}(1) \leq T = \sum_{i=1}^{r} \frac{p_i(1-p_i^k)(1-p_i^r)}{(1-p_i)^2} \leq kr$$

Corollary 1

To estimate $\mathrm{Tr}(\rho^i)$ for all $i \leq k$ within an additive error of ϵ and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$, it is necessary to estimate each $\mathrm{Tr}(\rho^j)$ for $j \leq r$ within an additive error of ϵ_j, as defined in **Theorem 1.** This can be achieved by using

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

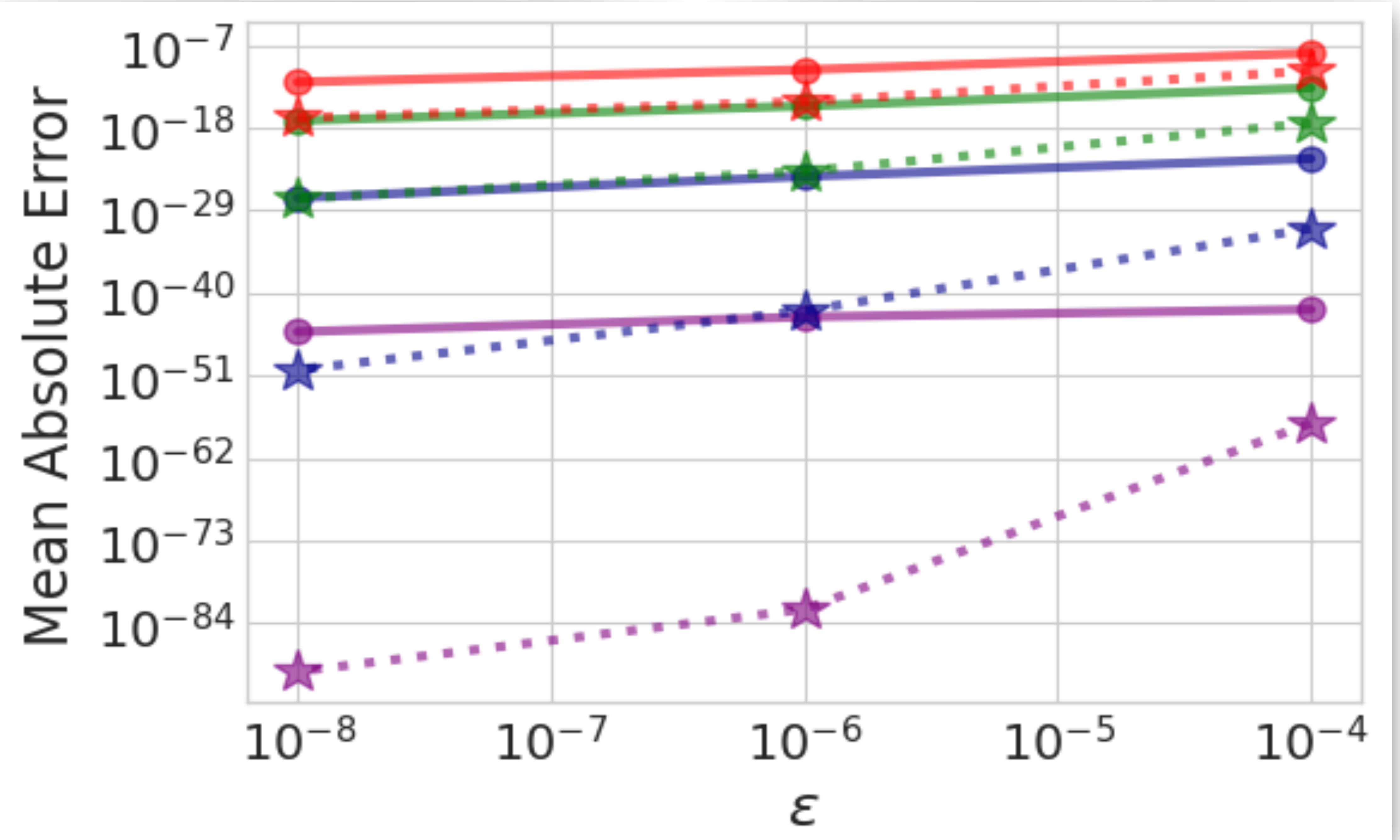
runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j)$ qubits and $\mathcal{O}(j)$ CSWAP operations. Here, T is defined in **Theorem 1.**

The value of T is significantly less than the upper bound in most cases.

Especially when the state is mixed (e.g., the largest eigenvalue is small), the system is large (e.g., r is large), or k grows large, the expected advantage becomes more dramatic compared to the upper bound.

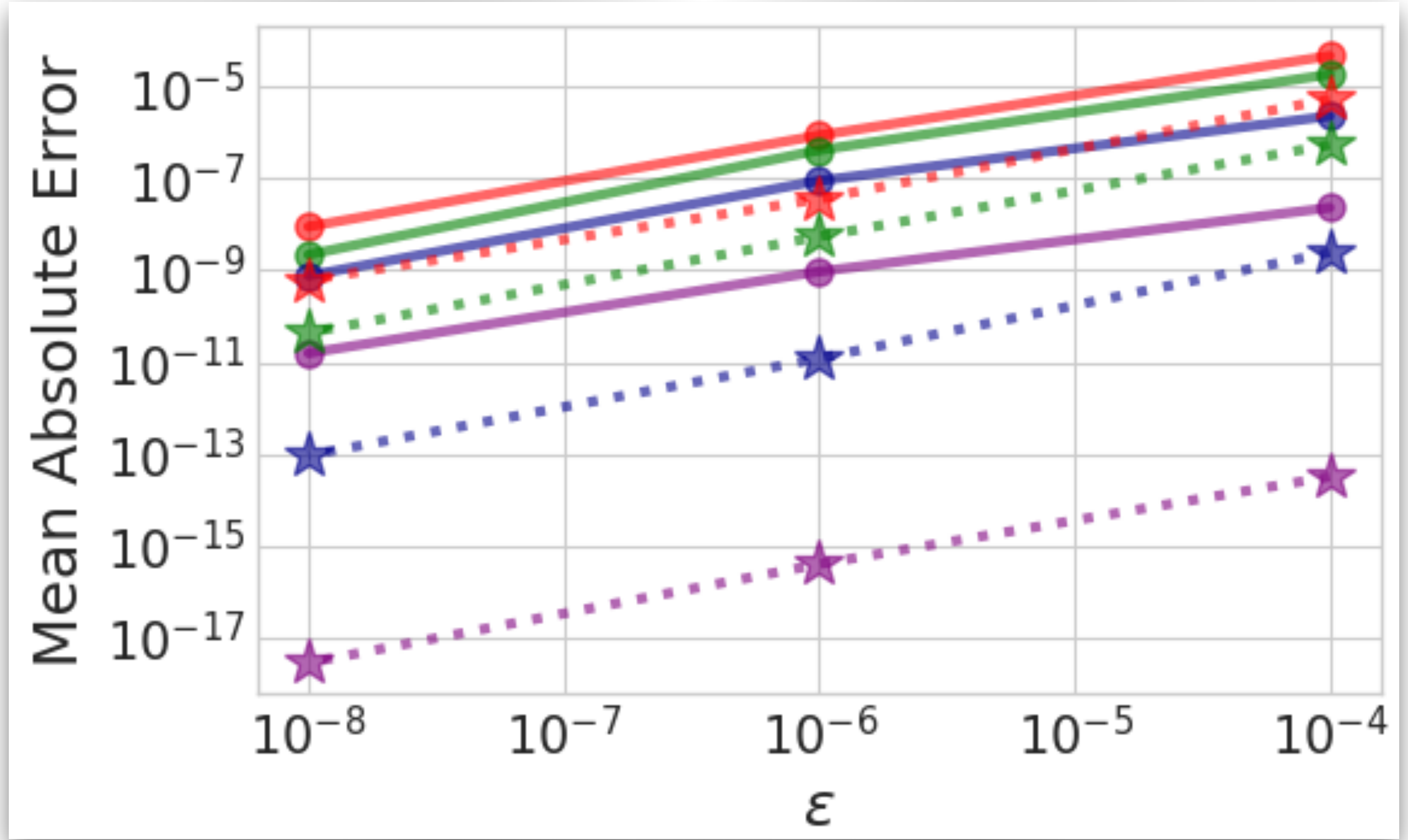
Numerical simulation

$$(r, k) = \left\{ \begin{array}{llll} (4,16) & (4,64) & (8,16) & (8,64) \\ (4,32) & (4,128) & (8,32) & (8,128) \end{array} \right\}$$



Arithmetically decaying

$T \sim \mathcal{O}(1)$



Geometrically decaying

$T \sim \mathcal{O}(kr)$

Overview

- Trace of powers & Literature review
- Mathematical intuitions
- Main results: algorithm, lemmas, theorems, corollaries
- Numerical simulations
- **Applications**
- Concluding remarks

Applications

Nonlinear functions

Computation of the **partition function** and other **thermodynamical** variables for the systems with finite energy levels and finite # of non interacting particles

e.g.

$$g(x) = e^{\beta x},\ x \in \mathbb{R}$$

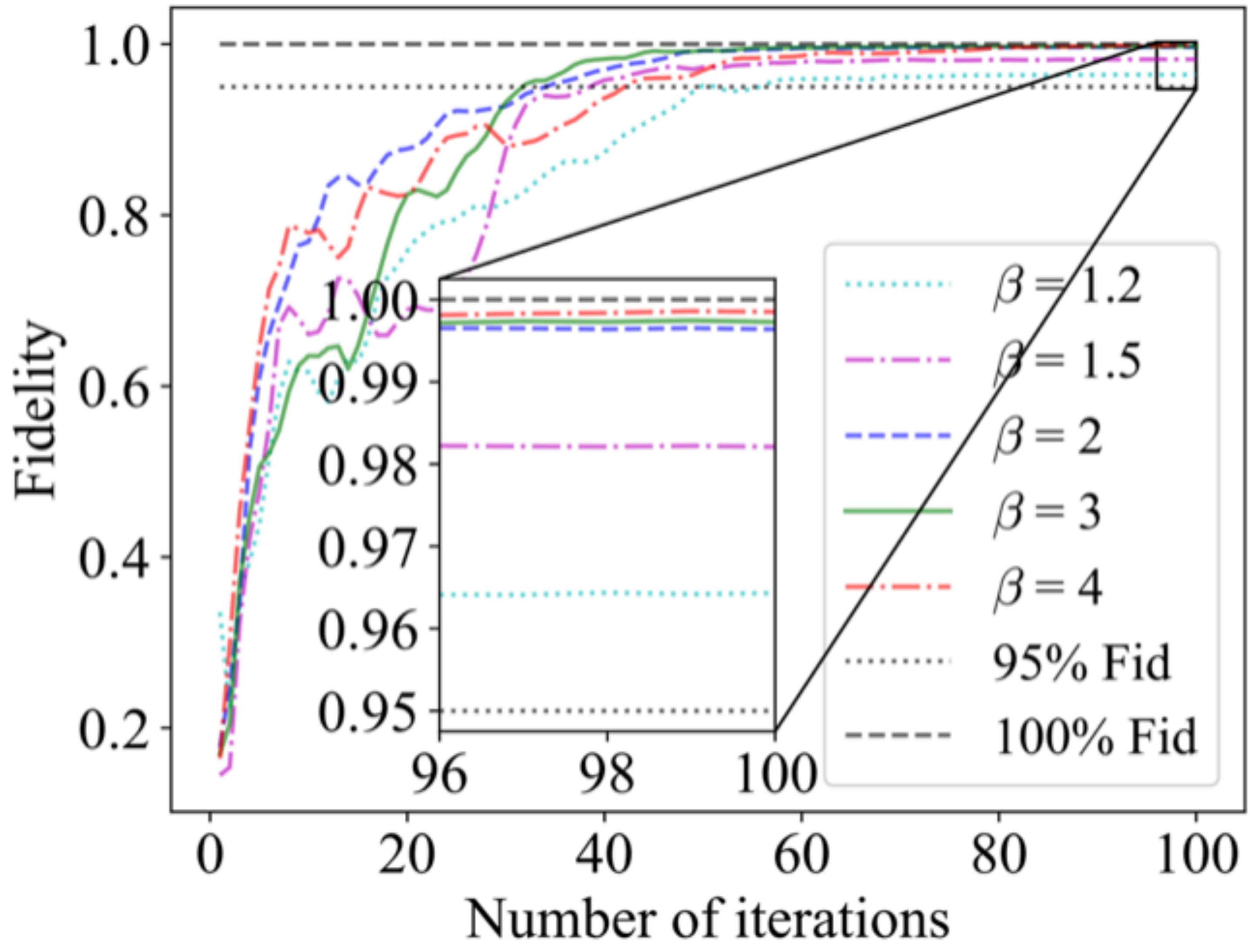
$$g(x) = (1+x)^{\alpha},\ \alpha \in \mathbb{R}^{+}$$

and more

Quantum Gibbs states

Variational quantum Gibbs state preparation with a **truncated Taylor series**

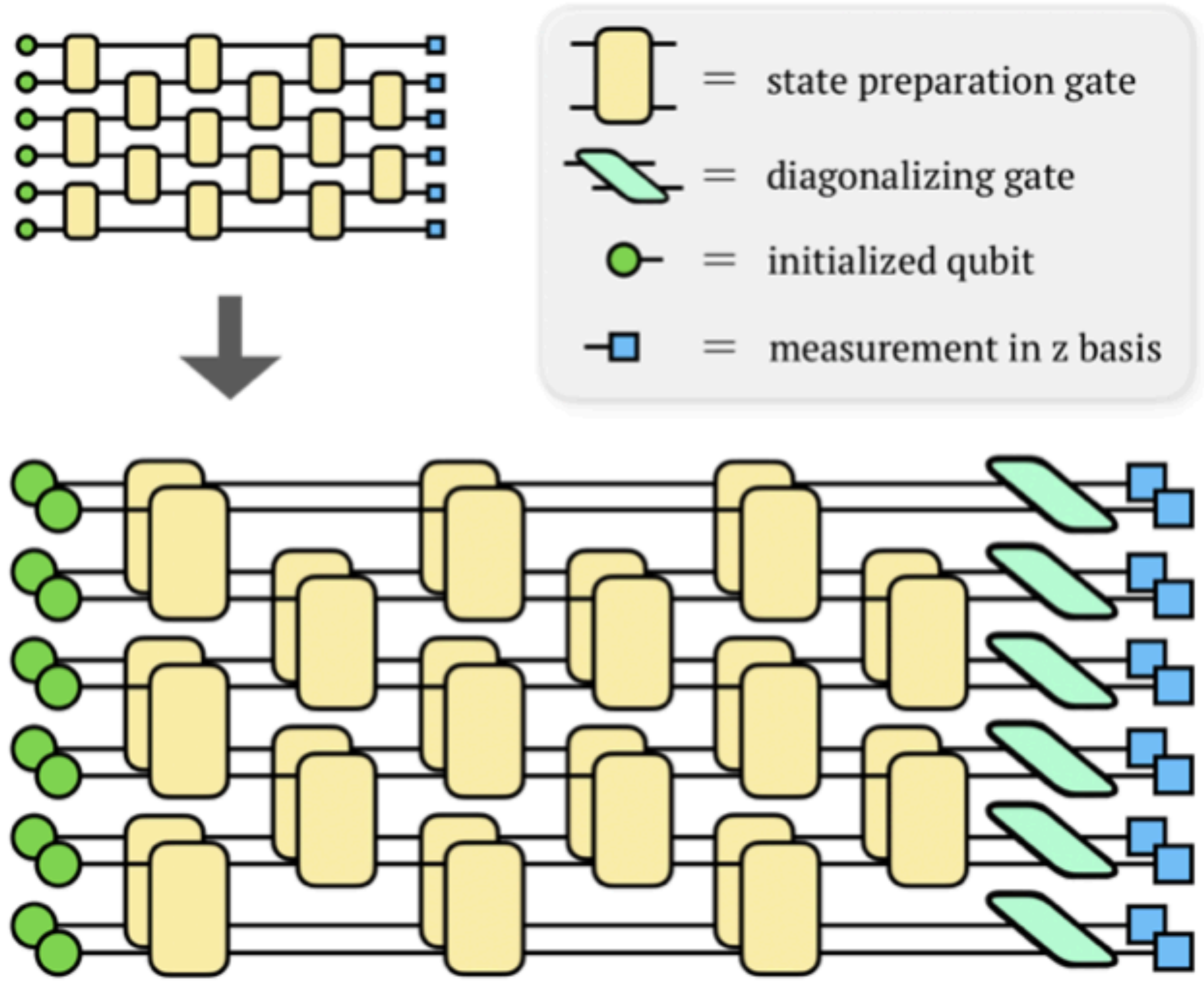
[Phys. Rev. Applied 16, 054035 (2021)]



Quantum error mitigation

Virtual distillation for quantum error mitigation

[Phys. Rev. X 11, 041036 (2021)]



Applications

Nonlinear functions

Computation of the **partition function** and other **thermodynamical** variables for the systems with finite energy levels and finite # of non interacting particles

e.g.

$$g(x) = e^{\beta x},\ x \in \mathbb{R}$$

$$g(x) = (1+x)^{\alpha},\ \alpha \in \mathbb{R}^+$$

and more

Quantum Gibbs states

Variational quantum Gibbs state preparation with a **truncated Taylor series**

[Phys. Rev. Applied 16, 054035 (2021)]

Quantum error mitigation

Virtual distillation for quantum error mitigation

[Phys. Rev. X 11, 041036 (2021)]

Calculating the nonlinear functions of quantum state

Let ρ be a quantum state with rank r.

Suppose there exist $\epsilon > 0$ and a slowly-growing function C (as a function of m) such that $g : \mathbb{R} \to \mathbb{R}$ is approximated by a degree m polynomial $f(x) = \sum_{k=0}^{m} c_k x^k$ on the interval $[0,1]$, in the sense that $\sup_{x \in [0,1]} |g(x) - f(x)| < \frac{\epsilon}{2r}$, and $\sum_{k=0}^{m} |c_k| < C$.

Calculating the nonlinear functions of quantum state

Then estimating $\mathrm{Tr}(g(\rho))$ within an ϵ additive error and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$ requires

Calculating the nonlinear functions of quantum state

Then estimating $\mathrm{Tr}(g(\rho))$ within an ϵ additive error and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$ requires

$$\mathcal{O}\left(\frac{T^2C^2r^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$ copies of ρ and

Calculating the nonlinear functions of quantum state

Then estimating $\mathrm{Tr}(g(\rho))$ within an ϵ additive error and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$ requires

$\mathcal{O}\left(\frac{T^2C^2r^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$ copies of ρ and

$\mathcal{O}\left(\frac{T^2C^2r}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$ runs on a constant-depth quantum circuit

Calculating the nonlinear functions of quantum state

Then estimating Tr$(g(\rho))$ within an ϵ additive error and

with a success probability of at least $1-\delta$, where $\delta \in (0,1)$ requires

$\mathcal{O}\left(\frac{T^2C^2r^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$ copies of ρ and

$\mathcal{O}\left(\frac{T^2C^2r}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$ runs on a constant-depth quantum circuit

consisting of $\mathcal{O}(r)$ qubits and $\mathcal{O}(r)$ CSWAP operations.

Calculating the nonlinear functions of quantum state

Previous: $\mathcal{O}\left(\frac{C^2m^2}{\epsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ copies of ρ and $\mathcal{O}\left(\frac{C^2m}{\epsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ runs on a circuit consisted of $\mathcal{O}(1)$ depth, $\mathcal{O}(m)$ qubits and $\mathcal{O}(m)$ CSWAP operations.

Calculating the nonlinear functions of quantum state

Previous: $\mathcal{O}\left(\frac{C^2m^2}{\epsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ copies of ρ and $\mathcal{O}\left(\frac{C^2m}{\epsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ runs on a circuit consisted of $\mathcal{O}(1)$ depth, $\mathcal{O}(m)$ qubits and $\mathcal{O}(m)$ CSWAP operations.

Ours: $\mathcal{O}\left(\frac{T^2C^2r^2}{\epsilon^2}\ln^2 r\ln\left(\frac{1}{\delta}\right)\right)$ copies of ρ and $\mathcal{O}\left(\frac{T^2C^2r}{\epsilon^2}\ln^2 r\ln\left(\frac{1}{\delta}\right)\right)$ runs on a circuit consisted of $\mathcal{O}(1)$ depth, $\mathcal{O}(r)$ qubits and $\mathcal{O}(r)$ CSWAP operations.

Calculating the nonlinear functions of quantum state

Previous: $\mathcal{O}\left(\frac{C^2m^2}{\epsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ copies of ρ and $\mathcal{O}\left(\frac{C^2m}{\epsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ runs on a circuit consisted of $\mathcal{O}(1)$ depth, $\mathcal{O}(m)$ qubits and $\mathcal{O}(m)$ CSWAP operations.

Typically, $m \gg r$,

so our enhanced theorem offers advantages for estimating $g(\rho)$.

Ours: $\mathcal{O}\left(\frac{T^2C^2r^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$ copies of ρ and $\mathcal{O}\left(\frac{T^2C^2r}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$ runs on a circuit consisted of $\mathcal{O}(1)$ depth, $\mathcal{O}(r)$ qubits and $\mathcal{O}(r)$ CSWAP operations.

Calculating the nonlinear functions of quantum state

Previous: $\mathcal{O}\left(\frac{C^2m^2}{\epsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ copies of ρ and $\mathcal{O}\left(\frac{C^2m}{\epsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ runs on a circuit consisted of $\mathcal{O}(1)$ depth, $\mathcal{O}(m)$ qubits and $\mathcal{O}(m)$ CSWAP operations.

Typically, $m \gg r$,

so our enhanced theorem offers advantages for estimating $g(\rho)$.

Ours: $\mathcal{O}\left(\frac{T^2C^2r^2}{\epsilon^2}\ln^2 r\ln\left(\frac{1}{\delta}\right)\right)$ copies of ρ and $\mathcal{O}\left(\frac{T^2C^2r}{\epsilon^2}\ln^2 r\ln\left(\frac{1}{\delta}\right)\right)$ runs on a circuit consisted of $\mathcal{O}(1)$ depth, $\mathcal{O}(r)$ qubits and $\mathcal{O}(r)$ CSWAP operations.

When $g(x) = e^{\beta x}$, C becomes $e^{|\beta|}$. We can efficiently estimate $\mathrm{Tr}(e^{\beta\rho})$ which has applications in thermodynamics and the density exponentiation algorithm.

Applications

Nonlinear functions

Computation of the **partition function** and other **thermodynamical** variables for the systems with finite energy levels and finite # of non interacting particles

e.g.

$$g(x) = e^{\beta x},\ x \in \mathbb{R}$$

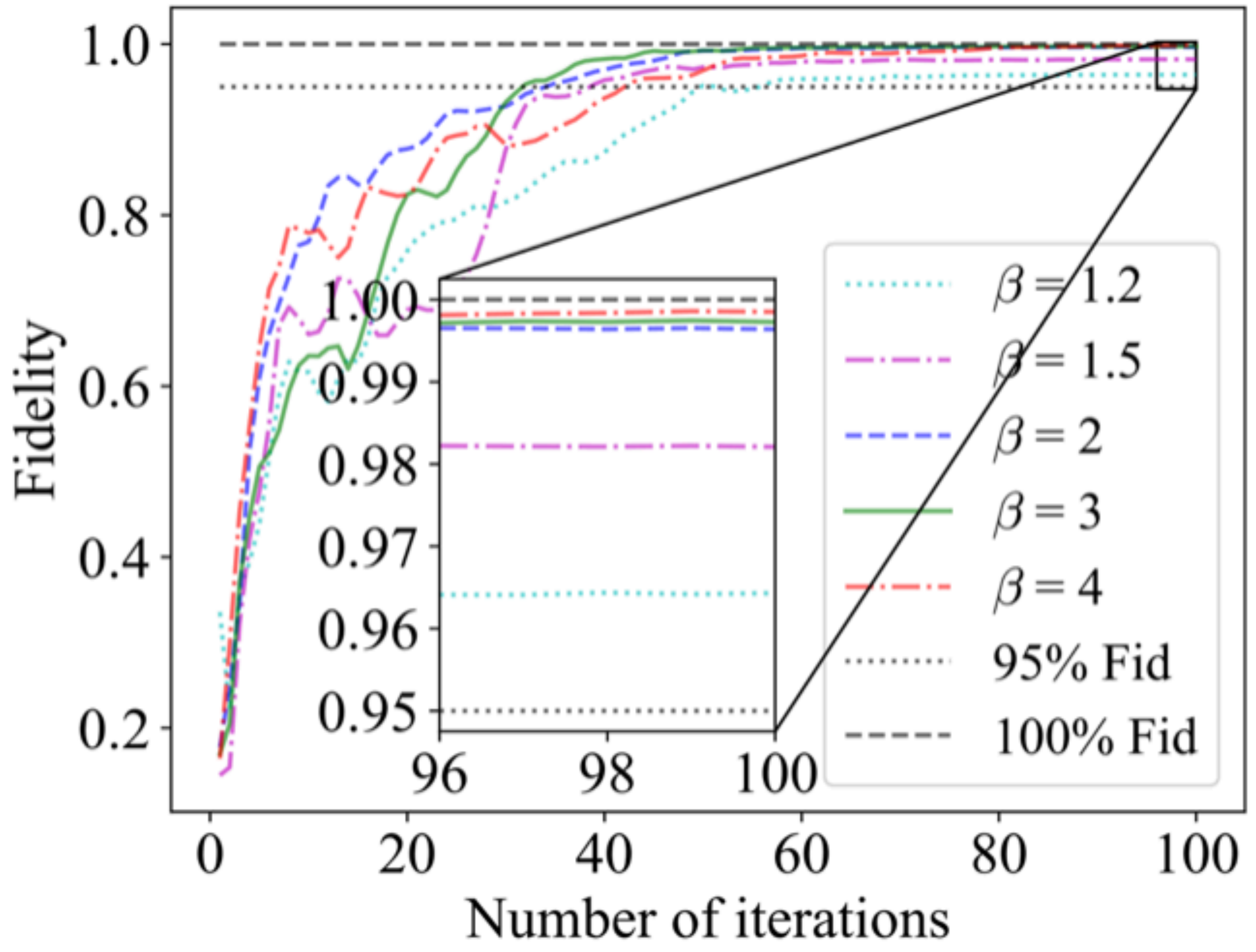
$$g(x) = (1+x)^{\alpha},\ \alpha \in \mathbb{R}^{+}$$

and more

Quantum Gibbs states

Variational quantum Gibbs state preparation with a **truncated Taylor series**

[Phys. Rev. Applied 16, 054035 (2021)]



Quantum error mitigation

Virtual distillation for quantum error mitigation

[Phys. Rev. X 11, 041036 (2021)]

Quantum Gibbs state preparation

The truncated Taylor series $S_k(\rho) = \sum_{i=1}^{k} \mathrm{Tr}\left(\left(\rho - I\right)^k \rho\right)$

is used as the **cost function for variational quantum Gibbs state preparation.**

Quantum Gibbs state preparation

The truncated Taylor series $S_k(\rho) = \sum_{i=1}^{k} \mathrm{Tr}\left(\left(\rho - I\right)^k \rho\right)$

is used as the **cost function for variational quantum Gibbs state preparation.**

It is shown that the fidelity $F\left(\rho(\theta_0), \rho_G\right)$ between the optimized state $\rho(\theta_0)$ and the Gibbs state ρ_G is bounded by $F(\rho(\theta_0), \rho_G) \geq 1 - \sqrt{2\left(\beta\epsilon + \frac{2r}{k+1}(1-\Delta)^{k+1}\right)}$

where β is the inverse temperature of the system,

and Δ is a constant that satisfies $-\Delta \ln(\Delta) < \frac{1}{k+1}(1-\Delta)^{k+1}$.

Quantum Gibbs state preparation

By using the inequality $D(\rho(\theta_0), \rho_G) < \sqrt{1 - F\left(\rho\left(\theta_0\right), \rho_G\right)},$ to achieve

$T\left(\rho(\theta_0), \rho_G\right) < \epsilon$, we need to set $k = \mathcal{O}\left(\,\cdot\,\right)$, where T is the trace distance.

Quantum Gibbs state preparation

By using the inequality $D(\rho(\theta_0), \rho_G) < \sqrt{1 - F\left(\rho\left(\theta_0\right), \rho_G\right)}$, to achieve $T\left(\rho(\theta_0), \rho_G\right) < \epsilon$, we need to set $k = \mathcal{O}\left(\cdot\right)$, where T is the trace distance.

Previous: $k = \mathcal{O}\left(\frac{r}{\epsilon^4}\right)$ qubits and CSWAP operations are needed.

Quantum Gibbs state preparation

By using the inequality $D(\rho(\theta_0), \rho_G) < \sqrt{1 - F\left(\rho\left(\theta_0\right), \rho_G\right)}$, to achieve $T\left(\rho(\theta_0), \rho_G\right) < \epsilon$, we need to set $k = \mathcal{O}\left(\,\cdot\,\right)$, where T is the trace distance.

Previous: $k = \mathcal{O}\left(\frac{r}{\epsilon^4}\right)$ qubits and CSWAP operations are needed.

Ours: $k = \mathcal{O}\left(r\right)$ qubits and CSWAP operations are needed.

Quantum Gibbs state preparation

By using the inequality $D(\rho(\theta_0), \rho_G) < \sqrt{1 - F\left(\rho\left(\theta_0\right), \rho_G\right)}$, to achieve $T\left(\rho(\theta_0), \rho_G\right) < \epsilon$, we need to set $k = \mathcal{O}\left(\,\cdot\,\right)$, where T is the trace distance.

Previous: $k = \mathcal{O}\left(\frac{r}{\epsilon^4}\right)$ qubits and CSWAP operations are needed.

Significantly reduces the number of qubits and CSWAP operations.

* Independent of the desired error bound ϵ.

Ours: $k = \mathcal{O}\left(r\right)$ qubits and CSWAP operations are needed.

Applications

Nonlinear functions

Computation of the **partition function** and other **thermodynamical** variables for the systems with finite energy levels and finite # of non interacting particles

e.g.

$$g(x) = e^{\beta x},\ x \in \mathbb{R}$$

$$g(x) = (1+x)^{\alpha},\ \alpha \in \mathbb{R}^+$$

and more

Quantum Gibbs states

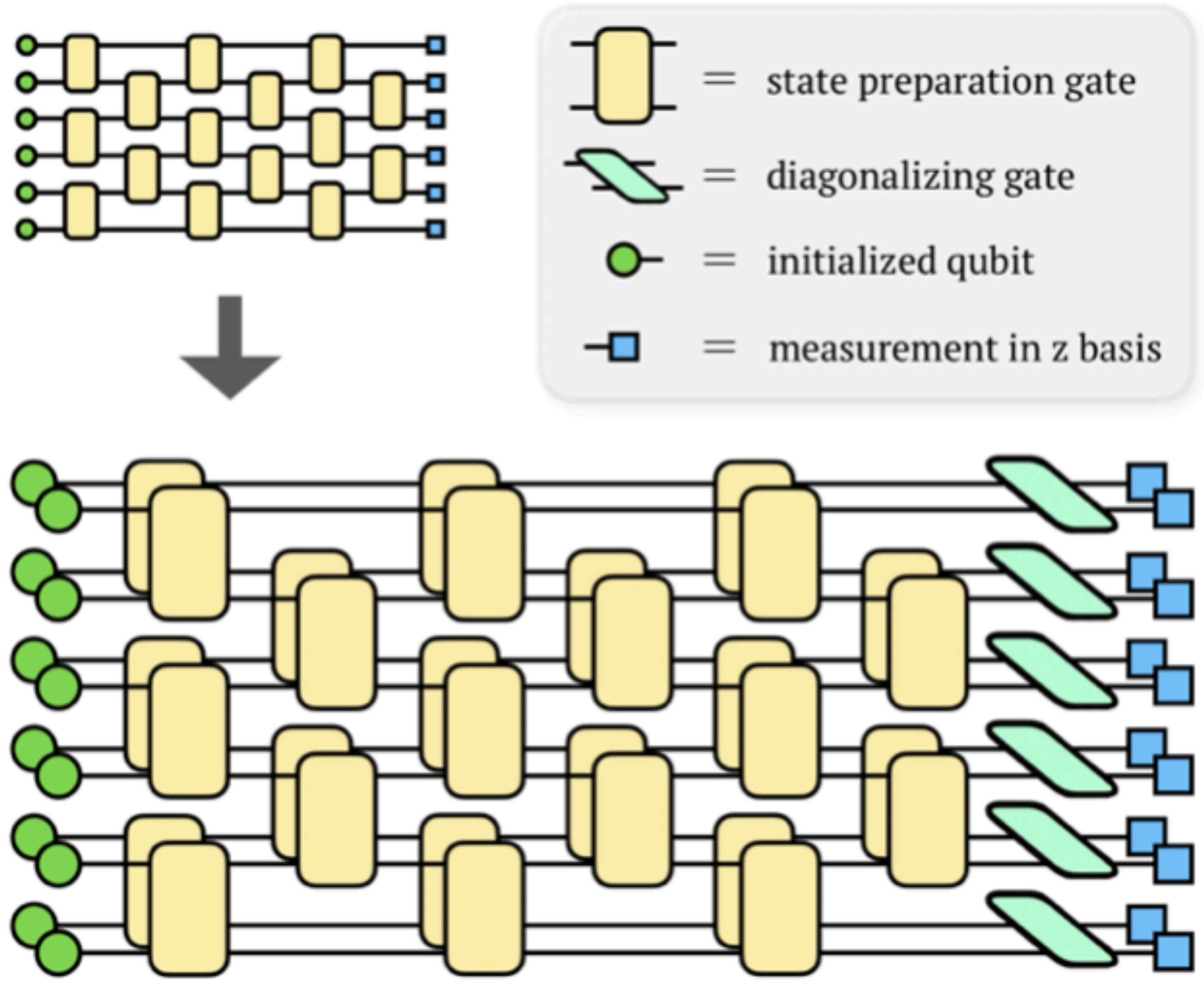
Variational quantum Gibbs state preparation with a **truncated Taylor series**

[Phys. Rev. Applied 16, 054035 (2021)]

Quantum error mitigation

Virtual distillation for quantum error mitigation

[Phys. Rev. X 11, 041036 (2021)]



Quantum error mitigation

The expected value of a Hermitian operator M is given by $\langle M \rangle = \mathrm{Tr}(M | \psi \rangle \langle \psi |)$.

Due to noise, this value becomes $\langle M \rangle_{\text{noise}} = \mathrm{Tr}(M\rho) \neq \langle M \rangle$.

Quantum error mitigation

The expected value of a Hermitian operator M is given by $\langle M \rangle = \mathrm{Tr}(M |\psi\rangle\langle\psi|)$.

Due to noise, this value becomes $\langle M \rangle_{\text{noise}} = \mathrm{Tr}(M\rho) \neq \langle M \rangle$.

Virtual distillation protocol offers a method to address this issue.

The protocol involves using collective measurements of k copies of the mixed state ρ to suppress incoherent errors.

Quantum error mitigation

The expected value of a Hermitian operator M is given by $\langle M \rangle = \mathrm{Tr}(M|\psi\rangle\langle\psi|)$.

Due to noise, this value becomes $\langle M \rangle_{\text{noise}} = \mathrm{Tr}(M\rho) \neq \langle M \rangle$.

Virtual distillation protocol offers a method to address this issue.

The protocol involves using collective measurements of k copies of the mixed state ρ to suppress incoherent errors.

This approach approximates the error-free expectation value $\langle M \rangle_{\text{vd}}^{(k)} = \dfrac{\mathrm{Tr}(M\rho^k)}{\mathrm{Tr}(\rho^k)}$,

where k denotes the number of copies used.

Quantum error mitigation

$$\langle M \rangle_{\mathrm{vd}}^{(k)} = \frac{\mathrm{Tr}(M\rho^k)}{\mathrm{Tr}(\rho^k)}$$

Quantum error mitigation

$$\langle M \rangle^{(k)}_{\mathrm{vd}} = \frac{\mathrm{Tr}(M\rho^k)}{\mathrm{Tr}(\rho^k)}$$

Corollary 1

To estimate $\mathrm{Tr}(\rho^i)$ for all $i \leq k$ within an additive error of ϵ and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$, it is necessary to estimate each $\mathrm{Tr}(\rho^j)$ for $j \leq r$ within an additive error of ε_j, as defined in **Theorem 1.** This can be achieved by using

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j)$ qubits and $\mathcal{O}(j)$ CSWAP operations. Here, T is defined in **Theorem 1.**

Quantum error mitigation

$$\langle M\rangle_{\mathrm{vd}}^{(k)} = \frac{\mathrm{Tr}(M\rho^k)}{\mathrm{Tr}(\rho^k)}$$

Corollary 2

Suppose there is an efficient decomposition $M = \sum_{\ell=1}^{N_M} a_\ell P_\ell$, where $a_\ell \in \mathbb{R}$ and $P_\ell = \sigma_{\ell_1} \otimes \cdots \otimes \sigma_{\ell_n}$ are tensor products of Pauli operators $\sigma_{\ell_1}, \ldots, \sigma_{\ell_n} \in \{\sigma_x, \sigma_y, \sigma_z, I\}$. The quantity $\sum_{\ell=1}^{N_M} |a_\ell| = \mathcal{O}(c)$ is bounded by a constant c.

To estimate Tr$(M\rho^i)$ for all $i \le k$ within an additive error of ϵ and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$, it is necessary to estimate each Tr$(M\rho^j)$ for $j \le r$ within an additive error of $\varepsilon_{j,M}$.

This can be achieved by using $\mathcal{O}\left(\frac{c^2 N_M}{\epsilon^2}\ln\left(\frac{1}{\delta}\right)\right)$ runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j)$ qubits and $\mathcal{O}(j)$ CSWAP operations, and estimating each Tr$(\rho^{j'})$ for $j' \le r$ within an additive error of $\varepsilon_{j'}$, by using $\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$ runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j')$ qubits and $\mathcal{O}(j')$ CSWAP operations. Here, $\varepsilon_{j,M}$, $\varepsilon_{j'}$ and T are defined in **Theorem 2.**

Corollary 1

To estimate Tr(ρ^i) for all $i \le k$ within an additive error of ϵ and with a success probability of at least $1-\delta$, where $\delta \in (0,1)$, it is necessary to estimate each Tr(ρ^j) for $j \le r$ within an additive error of ε_j, as defined in **Theorem 1.** This can be achieved by using

$$\mathcal{O}\left(\frac{T^2}{\epsilon^2}\ln^2 r \ln\left(\frac{1}{\delta}\right)\right)$$

runs on a constant-depth quantum circuit consisting of $\mathcal{O}(j)$ qubits and $\mathcal{O}(j)$ CSWAP operations. Here, T is defined in **Theorem 1.**

Overview

- Trace of powers & Literature review
- Mathematical intuitions
- Main results: algorithm, lemmas, theorems, corollaries
- Numerical simulations
- Applications
- **Concluding remarks**

Concluding remarks

- Our main contribution lies in proving that the **error increases linearly at most** when applying the Newton-Girard method with a recursive strategy.
- We also generalize the result to traces of powers with observables M, which are represented as $\mathrm{Tr}(M\rho^k)$.
- Our work can enhance any previous algorithms, including $\mathrm{Tr}(\rho^k)$ and/or $\mathrm{Tr}(M\rho^k)$.
- We can estimate the trace of powers with $\mathcal{O}(1)$-depth, $\mathcal{O}(r)$-width, and only $\mathcal{O}(r)$ -CSWAP operations.
- Our method also provides advantages in copy complexity when estimating the trace of large powers **with low-rank states or sufficiently mixed states.**

Future work

- It remains open for future work to find more applications that can take advantage of our work.
- Generalizing this result to multivariate trace estimation, or even Tr$(\rho^k\sigma^l)$, can open up more possibilities, such as calculating functions that satisfy the data-processing inequality under unital quantum channels, which can be an alternative tool for distance measures.
- Tightening the bounds on Theorems 1 and 2 is an interesting future research topic.
- **More about the effectiveness of our *"rank is all you need"* scheme.**

Thank you for listening!

Junseo Lee (harris.junseo@gmail.com)

arXiv:2408.00314