Heisenberg-limited ground state energy estimation

Yu Tong

Department of Mathematics, University of California, Berkeley

Joint work with Lin Lin (UC Berkeley, LBNL)

IPAM Quantum Numerical Linear Algebra Workshop, 2022

arXiv:2102.11340

Useful quantum advantage for quantum chemistry

- $H = \sum_{ij} h_{ij} c_i^\dagger c_j + \frac{1}{2} \sum_{ijkl} V_{ijkl} c_i^\dagger c_j^\dagger c_k c_l$;

Useful quantum advantage for quantum chemistry

- $H = \sum_{ij} h_{ij} c_i^\dagger c_j + \frac{1}{2} \sum_{ijkl} V_{ijkl} c_i^\dagger c_j^\dagger c_k c_l$;
- Extensive work has been done to perform Hamiltonian simulation e^{-iHt} since [Lloyd 1996]:
  - Trotter, Taylor series, Qubitization, QSP, commutator scaling, randomized, etc.

Useful quantum advantage for quantum chemistry

- $H = \sum_{ij} h_{ij} c_i^\dagger c_j + \frac{1}{2} \sum_{ijkl} V_{ijkl} c_i^\dagger c_j^\dagger c_k c_l$;
- Extensive work has been done to perform Hamiltonian simulation e^{-iHt} since [Lloyd 1996]:
  - Trotter, Taylor series, Qubitization, QSP, commutator scaling, randomized, etc.
- **Classical output:** we are looking for a classical output that can be generated from a quantum computer that is **useful** and at the same time **difficult** for classical computers.

Useful quantum advantage for quantum chemistry

- **Ground state energy:** useful in predicting material structures, simulating chemical reactions, etc.

Useful quantum advantage for quantum chemistry

- **Ground state energy:** useful in predicting material structures, simulating chemical reactions, etc.
- No provably efficient classical algorithm to attain chemical accuracy for strongly-correlated systems.

Useful quantum advantage for quantum chemistry

- **Ground state energy:** useful in predicting material structures, simulating chemical reactions, etc.
- No provably efficient classical algorithm to attain chemical accuracy for strongly-correlated systems.
- Without additional assumptions this problem is QMA-hard.

Assumption: a good initial guess

- We assume we can efficiently (in polynomial time) prepare a state ρ such that $\mathrm{Tr}[\rho\Pi_0] \geq \eta$ for some known η, here Π_0 is the ground state projection operator of H;

[1] Tubman, Mejuto-Zaera, Epstein, et al., 2018, arXiv:1809.05523.

Assumption: a good initial guess

- We assume we can efficiently (in polynomial time) prepare a state ρ such that $\mathrm{Tr}[\rho\Pi_0] \geq \eta$ for some known η, here Π_0 is the ground state projection operator of H;
- We say ρ is a **good initial guess** if η is "large";

[1] Tubman, Mejuto-Zaera, Epstein, et al., 2018, arXiv:1809.05523.

Assumption: a good initial guess

- We assume we can efficiently (in polynomial time) prepare a state ρ such that $\mathrm{Tr}[\rho\Pi_0] \geq \eta$ for some known η, here Π_0 is the ground state projection operator of H;
- We say ρ is a **good initial guess** if η is "large";
- Candidates for this ρ: Hartree-Fock state, VQE/QAOA.

[1] Tubman, Mejuto-Zaera, Epstein, et al., 2018, arXiv:1809.05523.

Assumption: a good initial guess

- We assume we can efficiently (in polynomial time) prepare a state ρ such that $\mathrm{Tr}[\rho\Pi_0] \geq \eta$ for some known η, here Π_0 is the ground state projection operator of H;
- We say ρ is a **good initial guess** if η is "large";
- Candidates for this ρ: Hartree-Fock state, VQE/QAOA.
- Caveat: we do not have a guarantee for the overlap, but empirically for medium sized quantum chemistry systems the overlap is usually acceptable. [1]

[1] Tubman, Mejuto-Zaera, Epstein, et al., 2018, arXiv:1809.05523.

Hardness with a good initial guess

- We consider the case where $\eta = 1/\mathsf{poly}(N)$;

Hardness with a good initial guess

- We consider the case where $\eta = 1/\mathrm{poly}(N)$;
- The assumption of a good initial guess makes the problem easier on the quantum side: it is in BQP for inverse polynomial precision.

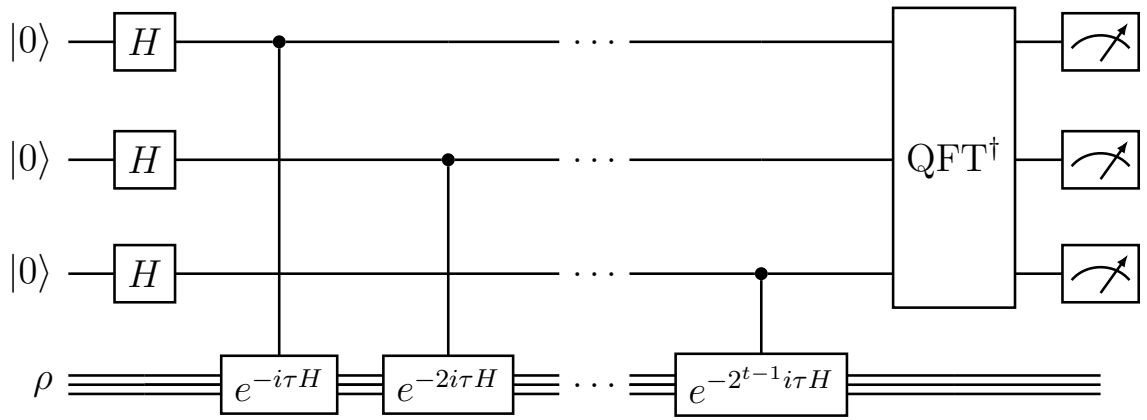
Hardness with a good initial guess

- We consider the case where $\eta = 1/\mathsf{poly}(N)$;
- The assumption of a good initial guess makes the problem easier on the quantum side: it is in BQP for inverse polynomial precision.
- However the problem is shown to be BQP-complete [Gharibian and Le Gall, 2021, arXiv:2111.09079].

Hardness with a good initial guess

- We consider the case where $\eta = 1/\text{poly}(N)$;
- The assumption of a good initial guess makes the problem easier on the quantum side: it is in BQP for inverse polynomial precision.
- However the problem is shown to be BQP-complete [Gharibian and Le Gall, 2021, arXiv:2111.09079].
- This means if a classical computer, given a good initial guess, can estimate the ground state energy to inverse polynomial precision in polynomial time, then $\mathrm{BQP} = \mathrm{BPP}$.

Quantum phase estimation



- **Heisenberg scaling** (total runtime $\mathcal{O}(\eta^{-2}\epsilon^{-1})$, ϵ is the ground state energy precision)
- Coherent time evolution $\mathcal{O}(\eta^{-1}\epsilon^{-1})$ (needs to maintain coherence, circuit depth).
- Requires multiple control qubits;

Three goals of our work

We propose an algorithm with the following three goals in mind:

(1) Heisenberg-limited precision scaling

(2) Using at most one ancilla qubit

(3) Using lower-depth circuit ($\mathcal{O}(\epsilon^{-1}\text{polylog}(\epsilon^{-1}\eta^{-1}))$)

Three goals of our work

We propose an algorithm with the following three goals in mind:

(1) Heisenberg-limited precision scaling

(2) Using at most one ancilla qubit

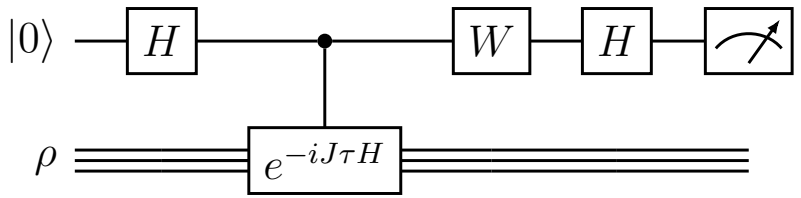
(3) Using lower-depth circuit ($\mathcal{O}(\epsilon^{-1}\mathrm{polylog}(\epsilon^{-1}\eta^{-1}))$)

These are desirable features for implementation on an early fault-tolerant quantum computer.

Algorithms	Requirements			Other issues
	(1)	(2)	(3)	
QPE (textbook version) [22, 48]	✓	✗	✗	
QPE (high-confidence) [37, 47, 56]	✓	✗	✓	
QPE (semi-classical QFT) [9, 30]	✓	✓	✗	
QPE (iterative) [34]	✓	✓	✓	Needs exact eigenstate ($p_0 = 1$)
The LCU approach [24]	✗	✗	✗	
The binary search approach [39]	✓	✗	✗	
VQE [45, 52, 54]	✗	✓	?	No precision guarantee
QITE [46]	✗	✓	?	Requires state tomography
QEEA [62]	✗	✓	✓	
Krylov subspace methods [31, 53, 63]	✗	✓	?	No precision guarantee
This work	✓	✓	✓	

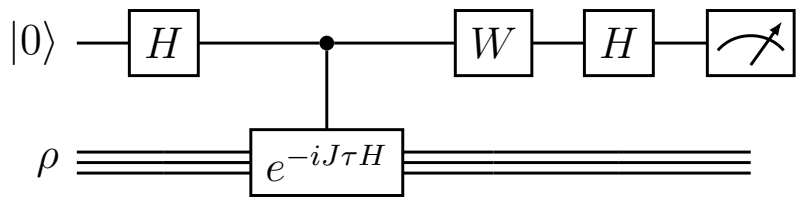
Figure: Comparison with other ground state energy estimation algorithms. The requirements are (1) achieving the Heisenberg-limited precision scaling, (2) using at most one ancilla qubit, and (3) the maximal evolution time being at most $\mathcal{O}(\epsilon^{-1}\text{polylog}(\epsilon^{-1}p_0^{-1}))$.

A simplified circuit



- Used in Kitaev's algorithm (iterative QPE).

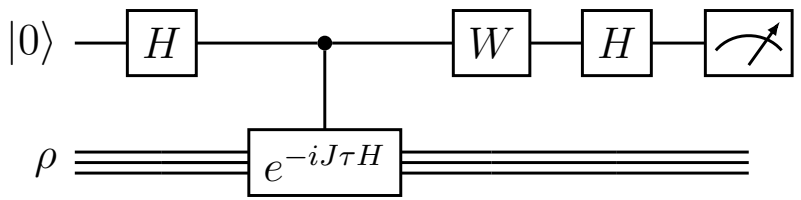
A simplified circuit



- Used in Kitaev's algorithm (iterative QPE).
- Denote the measurement outcome (± 1) by X (for $W = I$) and Y (for $W = S^\dagger$)

$$\mathbb{E}[X|J] = \mathrm{Re}\,\mathrm{Tr}[\rho e^{-iJ\tau H}]$$
$$\mathbb{E}[Y|J] = \mathrm{Im}\,\mathrm{Tr}[\rho e^{-iJ\tau H}]$$

A simplified circuit



- Used in Kitaev's algorithm (iterative QPE).
- Denote the measurement outcome (± 1) by X (for $W = I$) and Y (for $W = S^\dagger$)

$$\mathbb{E}[X|J] = \mathrm{Re}\,\mathrm{Tr}[\rho e^{-iJ\tau H}]$$
$$\mathbb{E}[Y|J] = \mathrm{Im}\,\mathrm{Tr}[\rho e^{-iJ\tau H}]$$

- We could estimate $\mathrm{Tr}[\rho e^{-iJ\tau H}]$ for $J = 1, 2, \ldots, d$ and do statistical inference [Somma2019, arXiv:1907.11748] (we are not doing this).

Introducing additional randomness

- If we estimate all $\mathrm{Tr}[\rho e^{-iJ\tau H}]$ for $J = 1, 2, \ldots, d$ then the total evolution time is

$$\tau + 2\tau + \cdots + d\tau = \frac{(d+1)d}{2}\tau.$$

Introducing additional randomness

- If we estimate all $\mathrm{Tr}[\rho e^{-iJ\tau H}]$ for $J = 1, 2, \ldots, d$ then the total evolution time is

$$\tau + 2\tau + \cdots + d\tau = \frac{(d+1)d}{2}\tau.$$

- We need to choose $\tau d = \epsilon^{-1}$ and this results in $\mathcal{O}(\tau^{-1}\epsilon^{-2})$ total evolution time. Cannot saturate the Heisenberg limit.

Introducing additional randomness

- If we estimate all $\mathrm{Tr}[\rho e^{-iJ\tau H}]$ for $J = 1, 2, \ldots, d$ then the total evolution time is

$$\tau + 2\tau + \cdots + d\tau = \frac{(d+1)d}{2}\tau.$$

- We need to choose $\tau d = \epsilon^{-1}$ and this results in $\mathcal{O}(\tau^{-1}\epsilon^{-2})$ total evolution time. Cannot saturate the Heisenberg limit.
- **Idea:** sample J from a distribution use the idea of importance sampling.

Random evolution time

- From our previous definition of random variables X and Y we define $Z = X + iY \in \{\pm 1 \pm i\}$, which satisfies

$$\mathbb{E}[Z|J] = \mathrm{Tr}[\rho e^{-iJ\tau H}]$$

Random evolution time

- From our previous definition of random variables X and Y we define $Z = X + iY \in \{\pm 1 \pm i\}$, which satisfies

$$\mathbb{E}[Z|J] = \mathrm{Tr}[\rho e^{-iJ\tau H}]$$

- With random evolution time J we have

$$\begin{aligned}\mathbb{E}[Ze^{i(\theta_J + Jx)}] &= \sum_{j=-d}^{d} \mathbb{E}[Z|J=j]\,\mathrm{Pr}[J=j]e^{i(\theta_j + jx)} \\ &= \sum_{j=-d}^{d} \mathrm{Tr}[\rho e^{-ij\tau H}]\,\mathrm{Pr}[J=j]e^{i(\theta_j + jx)},\end{aligned}$$

and we view this as a function of $x \in \mathbb{R}$.

- For a function $f(w) = \sum_{j=-d}^{d} A_j e^{ijw}$, with $\sum |A_j| = 1$, we can choose $\Pr[J = j] = |A_j|$ and $\theta_j = \arg A_j$, then

$$\mathbb{E}[Ze^{i(\theta_J + Jx)}] = \mathrm{Tr}[\rho f(x - \tau H)].$$

- For a function $f(w) = \sum_{j=-d}^{d} A_j e^{ijw}$, with $\sum |A_j| = 1$, we can choose $\Pr[J = j] = |A_j|$ and $\theta_j = \arg A_j$, then

$$\mathbb{E}[Z e^{i(\theta_J + Jx)}] = \mathrm{Tr}[\rho f(x - \tau H)].$$

- The expectation value can be evaluated using Monte Carlo sampling.

- For a function $f(w) = \sum_{j=-d}^{d} A_j e^{ijw}$, with $\sum |A_j| = 1$, we can choose $\Pr[J = j] = |A_j|$ and $\theta_j = \arg A_j$, then

$$\mathbb{E}[Z e^{i(\theta_J + Jx)}] = \mathrm{Tr}[\rho f(x - \tau H)].$$

- The expectation value can be evaluated using Monte Carlo sampling.
- Do not need to re-sample for each x (can reuse Z)!

- For a function $f(w) = \sum_{j=-d}^{d} A_j e^{ijw}$, with $\sum |A_j| = 1$, we can choose $\Pr[J = j] = |A_j|$ and $\theta_j = \arg A_j$, then

$$\mathbb{E}[Z e^{i(\theta_J + Jx)}] = \mathrm{Tr}[\rho f(x - \tau H)].$$

- The expectation value can be evaluated using Monte Carlo sampling.
- Do not need to re-sample for each x (can reuse Z)!
- If a function f can be approximated by a finite Fourier sum then $\mathrm{Tr}[\rho f(x - \tau H)]$ can be approximately computed using this method, up to an approximation error.

The cumulative distribution function

- We consider a specific $f(w)$:

$$f(w) = \begin{cases} 1, & w \geq 0, \\ 0, & w < 0. \end{cases}$$

The cumulative distribution function

- We consider a specific $f(w)$:

$$f(w) = \begin{cases} 1, & w \geq 0, \\ 0, & w < 0. \end{cases}$$

- We write down the eigendecomposition of the Hamiltonian

$$H = \sum_k \lambda_k \Pi_k,$$

and denote $p_k = \mathrm{Tr}[\Pi_k \rho]$ which is the population of the initial state ρ in the k-th eigensubspace.

- Then the quantity we can estimate is

$$C(x) := \mathrm{Tr}[\rho f(x - \tau H)] = \sum_{k:\tau\lambda_k \leq x} p_k,$$

- Then the quantity we can estimate is

$$C(x) := \mathrm{Tr}[\rho f(x - \tau H)] = \sum_{k:\tau\lambda_k \leq x} p_k,$$

- This is the **cumulative distribution function (CDF)** of the spectral measure.

- Then the quantity we can estimate is

$$C(x) := \mathrm{Tr}[\rho f(x - \tau H)] = \sum_{k:\tau\lambda_k \leq x} p_k,$$

- This is the **cumulative distribution function (CDF)** of the spectral measure.
- Why should we study it: each jump corresponds to an eigenvalue!

- Then the quantity we can estimate is

$$C(x) := \mathrm{Tr}[\rho f(x - \tau H)] = \sum_{k:\tau\lambda_k \leq x} p_k,$$

- This is the **cumulative distribution function (CDF)** of the spectral measure.
- Why should we study it: each jump corresponds to an eigenvalue!
- Caveat: this is a discontinuous function.

The CDF: the numerical result

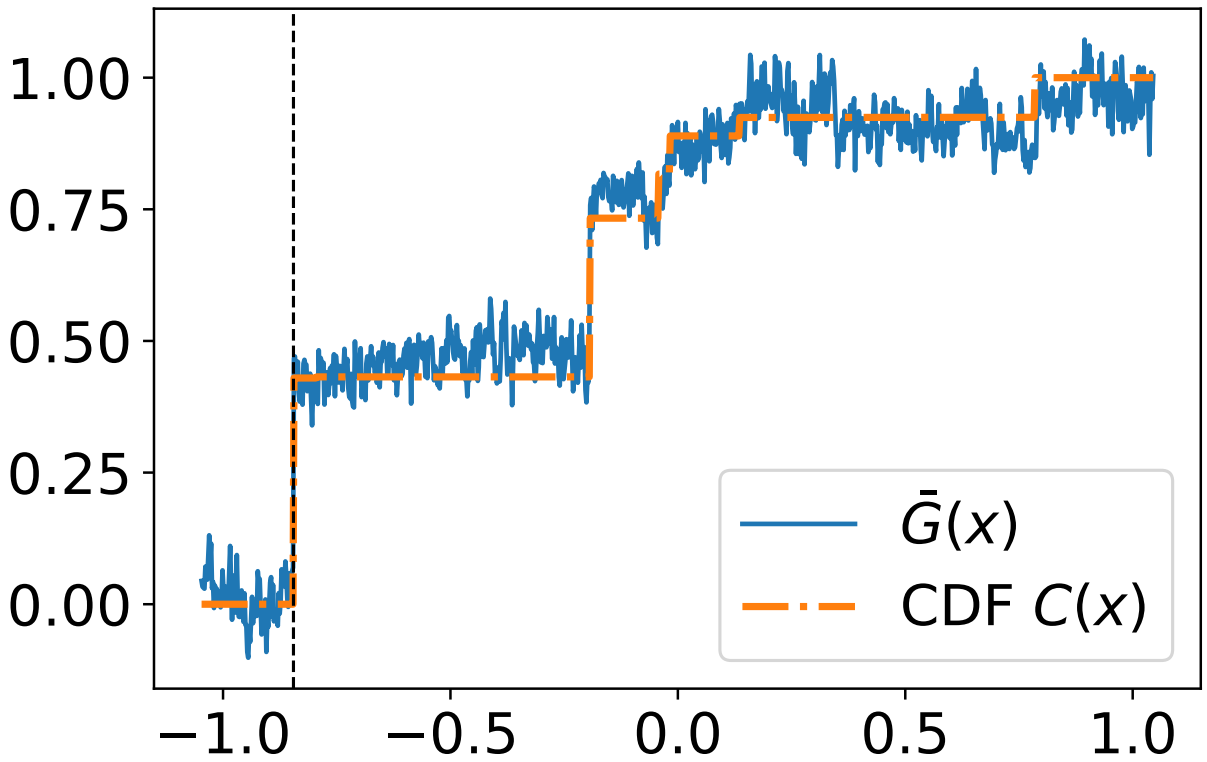


Figure: Evaluating the CDF by sampling from the quantum circuit. Note that we do not need to re-sample for each point.

The CDF: the numerical result

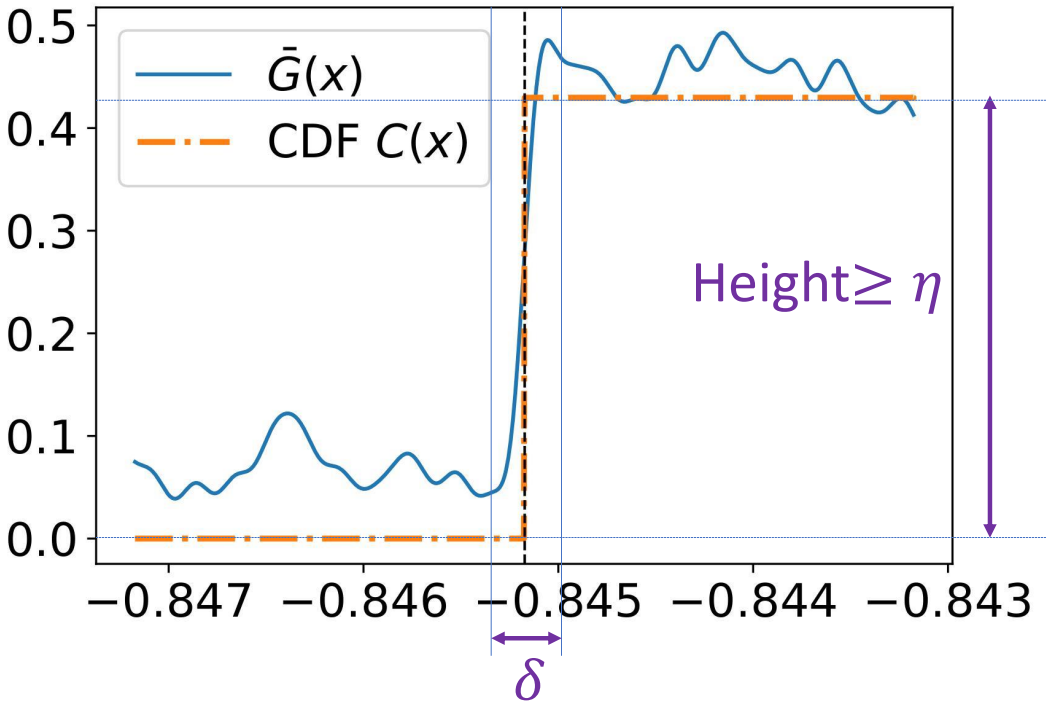


Figure: Zoom-in around the ground state energy

Summary of the algorithm

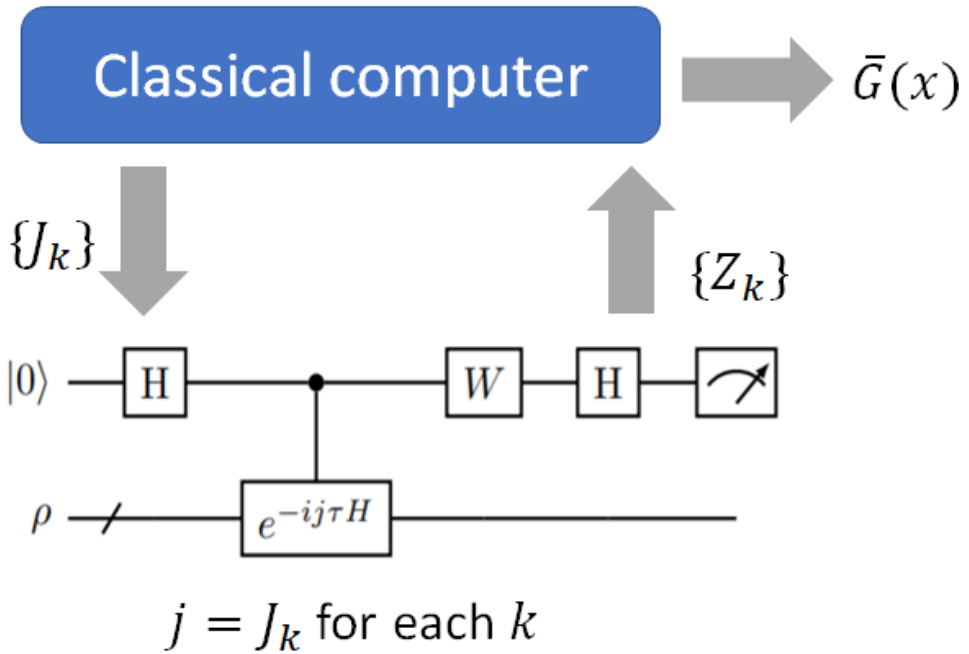


Figure: Schematic representation of the algorithm to compute the CDF. $\bar{G}(x)$ is the approximate CDF.

Ground state energy estimation

- Only need to find the first jump in the CDF; can be done using a binary search;

Ground state energy estimation

- Only need to find the first jump in the CDF; can be done using a binary search;
- The transition length is $\delta = \tau\epsilon$ to ensure precision ϵ;

Ground state energy estimation

- Only need to find the first jump in the CDF; can be done using a binary search;
- The transition length is $\delta = \tau\epsilon$ to ensure precision ϵ;
- The max (coherent) evolution time is $\tau/\delta = \mathcal{O}(\epsilon^{-1})$ (compare with $\mathcal{O}(\epsilon^{-1}\eta^{-1})$ for QPE, lower circuit depth);

Ground state energy estimation

- Only need to find the first jump in the CDF; can be done using a binary search;
- The transition length is $\delta = \tau\epsilon$ to ensure precision ϵ;
- The max (coherent) evolution time is $\tau/\delta = \mathcal{O}(\epsilon^{-1})$ (compare with $\mathcal{O}(\epsilon^{-1}\eta^{-1})$ for QPE, lower circuit depth);
- Monte Carlo accuracy $\mathcal{O}(\eta)$, therefore $N_s = \mathcal{O}(\eta^{-2})$;

Ground state energy estimation

- Only need to find the first jump in the CDF; can be done using a binary search;
- The transition length is $\delta = \tau\epsilon$ to ensure precision ϵ;
- The max (coherent) evolution time is $\tau/\delta = \mathcal{O}(\epsilon^{-1})$ (compare with $\mathcal{O}(\epsilon^{-1}\eta^{-1})$ for QPE, lower circuit depth);
- Monte Carlo accuracy $\mathcal{O}(\eta)$, therefore $N_s = \mathcal{O}(\eta^{-2})$;
- Total evolution time is $\mathcal{O}(N_s\tau d) = \mathcal{O}(\epsilon^{-1}\eta^{-2})$ **(Heisenberg scaling)**;

Comparison with QPE

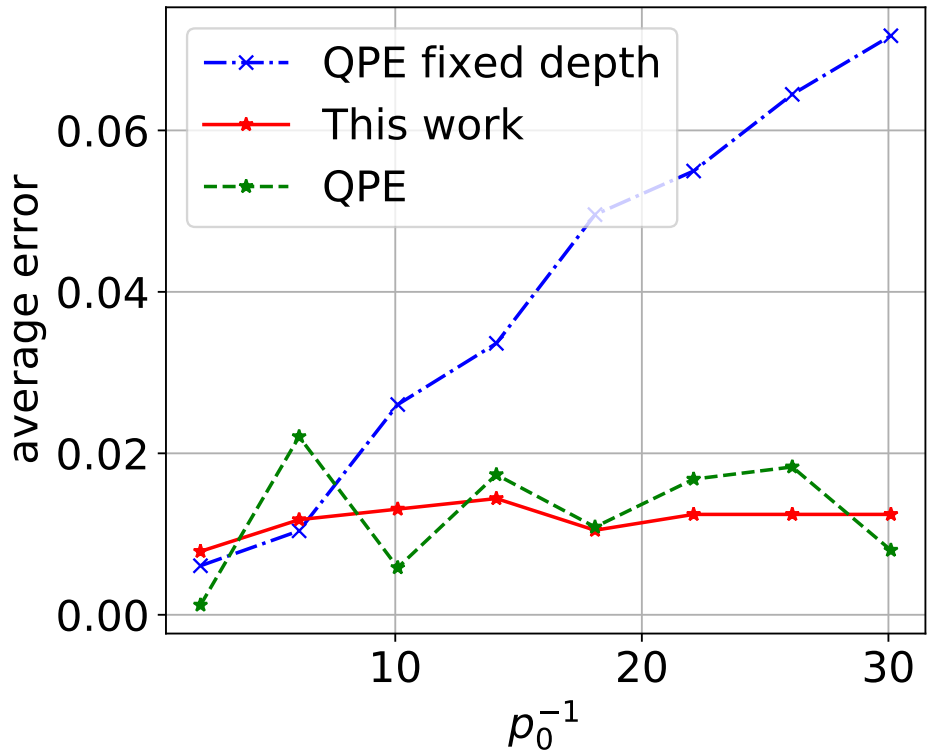


Figure: Average ground state energy estimation error for this method and QPE with fixed max evolution time, benchmarked against QPE with increasing max evolution time. $p_0 = \mathrm{Tr}[\rho\Pi_0]$.

Comparison with QPE

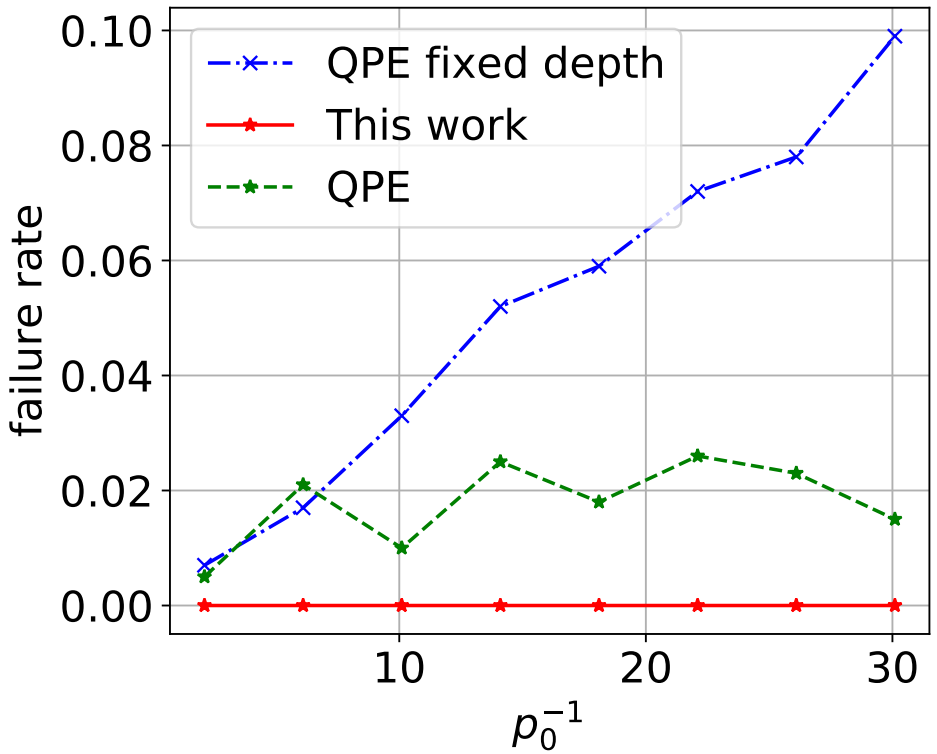


Figure: Failure (error beyond a certain threshold) rates

Fourier approximation

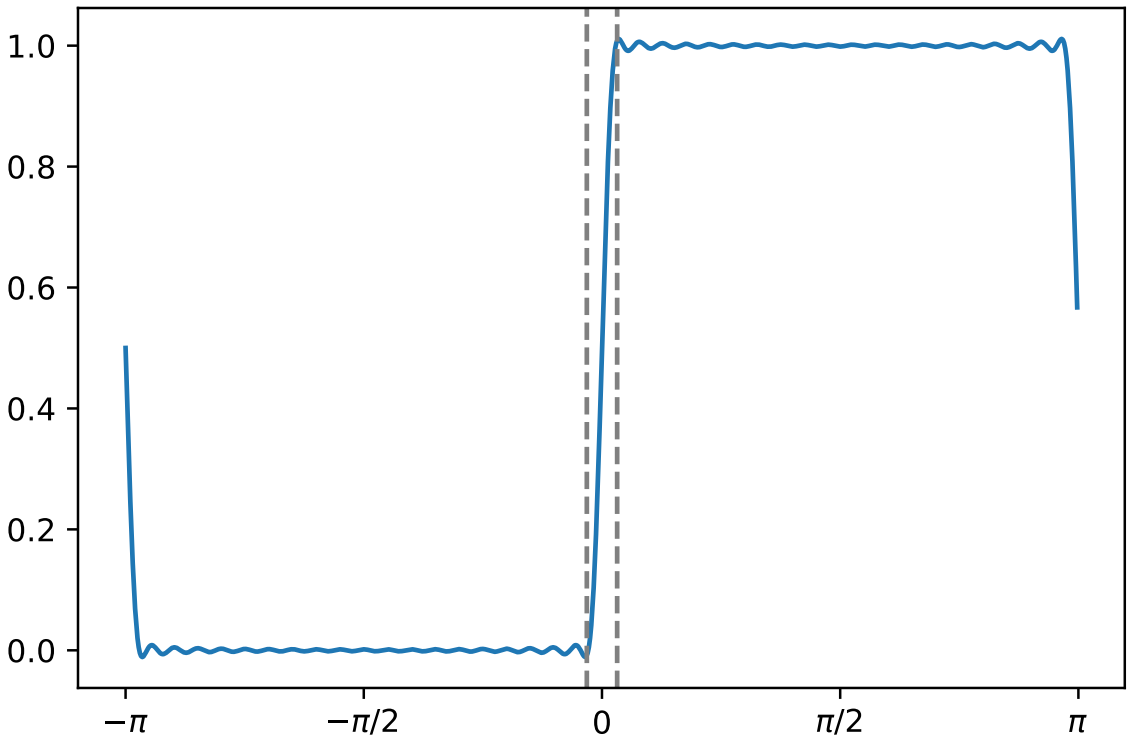


Figure: Our construction of the approx Heaviside function. The number of terms is $2d + 1$ where $d = \mathcal{O}(\delta^{-1} \log(\epsilon'^{-1}))$.

Fourier approximation

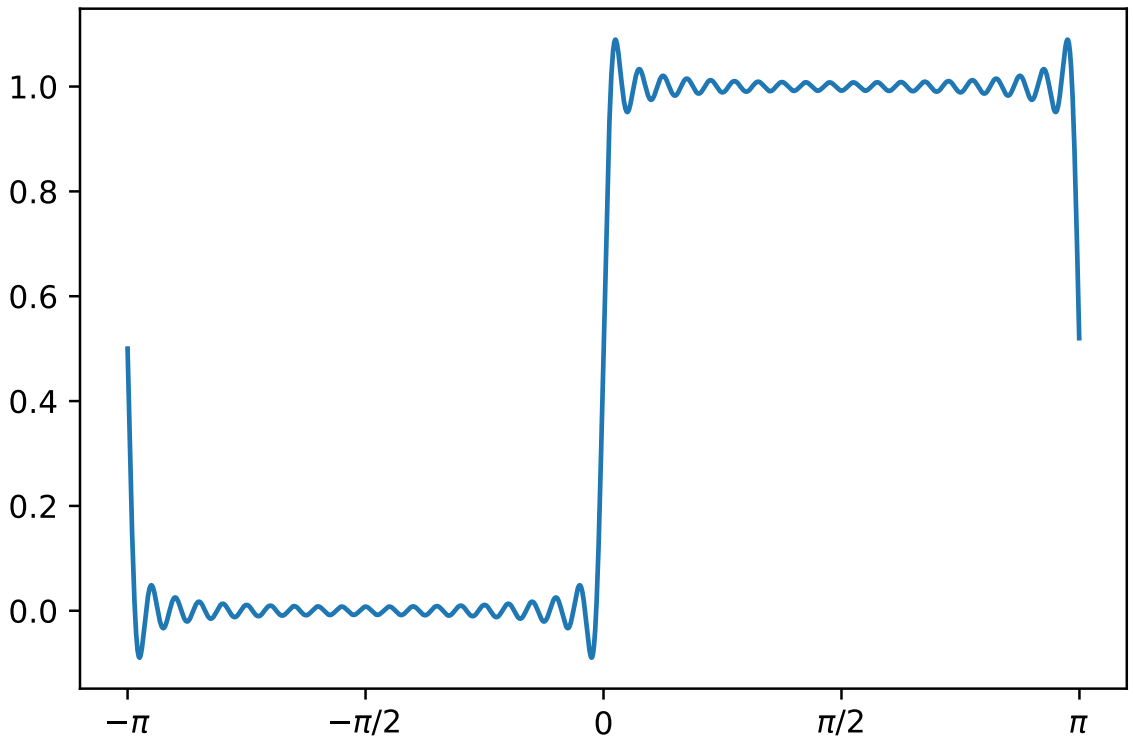
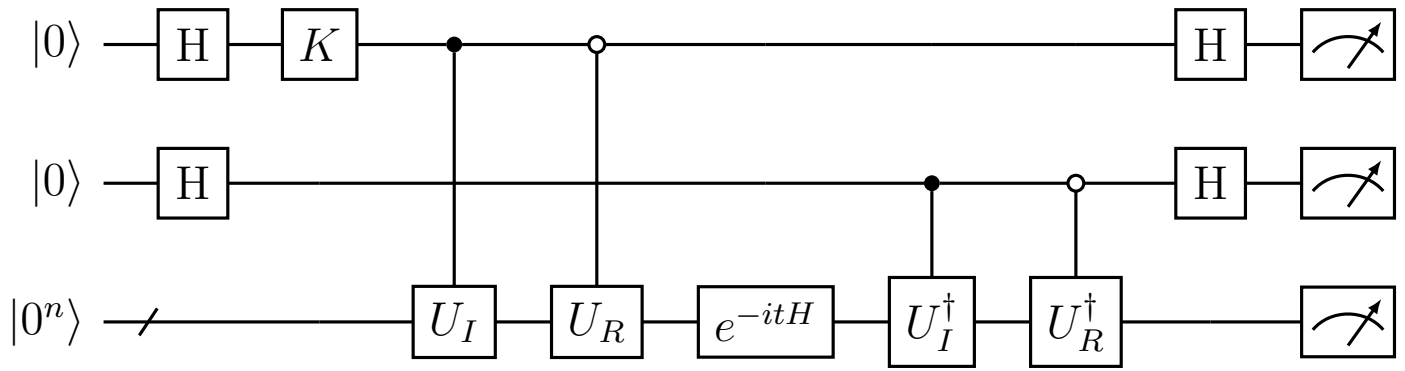


Figure: Directly truncating the Fourier expansion of the Heaviside function. Note the Gibbs phenomenon.

The control-free version



- We assume we have a reference eigenstate $|\psi_R\rangle$ that is easy to prepare and corresponds to a known eigenvalue.[2,3]
- Example: for system with particle number conservation we can use the vacuum state.

[2] Huggins, Lee, Baek, O'Gorman, Whaley, 2019, arXiv:1909.09114
[3] Russo, Rudinger, Morrison, Baczewski, 2020, arXiv:2007.08697

Observables and unbiased time evolution

- Estimating observable expectation values:

  Zhang, Wang, and Johnson, 2021, "Computing Ground State Properties with Early Fault-Tolerant Quantum Computers".

- Unbiased estimation of $\mathrm{Tr}[\rho e^{-iHt}]$ with cost independent of the number of terms:

  Wan, Berta, and Campbell, 2021, "A randomized quantum algorithm for statistical phase estimation".

Conclusions

- We developed an algorithm for ground state energy estimation that has the following features:
  1. Achieves Heisenberg scaling;
  2. Uses few ancilla qubits (one for controlled version, two for control-free version);
  3. Requires shorter circuit depth than QPE.

[4] Somma, 2019, arXiv:1907.11748

Conclusions

- We developed an algorithm for ground state energy estimation that has the following features:
  1. Achieves Heisenberg scaling;
  2. Uses few ancilla qubits (one for controlled version, two for control-free version);
  3. Requires shorter circuit depth than QPE.

- The CDF could be of use in itself.

[4] Somma, 2019, arXiv:1907.11748

Conclusions

- We developed an algorithm for ground state energy estimation that has the following features:
  1. Achieves Heisenberg scaling;
  2. Uses few ancilla qubits (one for controlled version, two for control-free version);
  3. Requires shorter circuit depth than QPE.

- The CDF could be of use in itself.

- A general method to evaluate $\mathrm{Tr}[\rho f(x - \tau H)]$ for given smooth 2π-periodic function f.[4]

[4] Somma, 2019, arXiv:1907.11748