The Heisenberg limit and early fault-tolerant quantum algorithms

Yu Tong

Institute for Quantum Information and Matter, Caltech

September, 2023
Practical quantum advantage in quantum chemistry

Can we solve a **practically useful** problem on a quantum computer faster than on a classical computer?

---

Can we solve a **practically useful** problem on a quantum computer faster than on a classical computer?

**Quantum chemistry** may be the right place to look.

---

Can we solve a practically useful problem on a quantum computer faster than on a classical computer?

Quantum chemistry may be the right place to look.

The basic problem: the ground state energy (lowest eigenvalue of $H$).

---

1. Bogojesk et al., 2020, *Quantum chemical accuracy from density functional approximations via machine learning*. 
Practical quantum advantage in quantum chemistry

- Can we solve a **practically useful** problem on a quantum computer faster than on a classical computer?

- **Quantum chemistry** may be the right place to look.

- The basic problem: the ground state energy (lowest eigenvalue of $H$).

- Compare with classical algorithms: need very high accuracy.
  - Density functional theory can get to precision of 2-3 kcal·mol$^{-1}$.
  - Chemical accuracy: 1 kcal·mol$^{-1}$.

- We should care **very much** about how the cost of the quantum algorithm scales with precision.

---

1 Bogojesk et al., 2020, *Quantum chemical accuracy from density functional approximations via machine learning.*
Example: from eigenstate to eigenvalue

- The precision scaling is usually more complicated in the quantum setting.
Example: from eigenstate to eigenvalue

- The precision scaling is usually more complicated in the quantum setting.

- For a Hamiltonian $H = \sum_{i=1}^{M} \alpha_i P_i$ ($P_i$ is a Pauli operator), given an eigenstate $|\Psi\rangle$, how to get the eigenvalue $\lambda$?

Classical computer: $\lambda = \langle \Psi | H | \Psi \rangle$ (one matrix-vector multiplication, one inner product, machine precision).

Quantum computer: measure each Pauli operator, obtain $0/1$ outputs, take average to get $\langle \Psi | P_i | \Psi \rangle$, then add up all Pauli terms.
Example: from eigenstate to eigenvalue

- The precision scaling is usually more complicated in the quantum setting.

- For a Hamiltonian $H = \sum_{i=1}^{M} \alpha_i P_i$ ($P_i$ is a Pauli operator), given an eigenstate $|\Psi\rangle$, how to get the eigenvalue $\lambda$?

- Classical computer: $\lambda = \langle \Psi | H | \Psi \rangle$ (one matrix-vector multiplication, one inner product, machine precision).
Example: from eigenstate to eigenvalue

- The precision scaling is usually more complicated in the quantum setting.

- For a Hamiltonian $H = \sum_{i=1}^{M} \alpha_i P_i$ ($P_i$ is a Pauli operator), given an eigenstate $|\Psi\rangle$, how to get the eigenvalue $\lambda$?

- Classical computer: $\lambda = \langle \Psi | H | \Psi \rangle$ (one matrix-vector multiplication, one inner product, machine precision).

- Quantum computer: measure each Pauli operator, obtain 0/1 outputs, take average to get $\langle \Psi | P_i | \Psi \rangle$, then add up all Pauli terms.
Example: to measure $X \otimes X$, we can apply $\text{Had} \otimes \text{Had}$ to the quantum state, so that we can now measure in the computational basis ($Z \otimes Z$):

$$
\langle \Psi | X \otimes X | \Psi \rangle = \langle \Psi | (\text{Had} \otimes \text{Had})(Z \otimes Z)(\text{Had} \otimes \text{Had})| \Psi \rangle.
$$
Example: to measure $X \otimes X$, we can apply $\text{Had} \otimes \text{Had}$ to the quantum state, so that we can now measure in the computational basis $(Z \otimes Z)$:

$$\langle \Psi | X \otimes X | \Psi \rangle = \langle \Psi | (\text{Had} \otimes \text{Had})(Z \otimes Z)(\text{Had} \otimes \text{Had}) | \Psi \rangle .$$

We will get a random output $\hat{m} \in \{0, 1\}$, such that

$$\mathbb{E}[(-1)^{\hat{m}}] = \langle \Psi | X \otimes X | \Psi \rangle .$$
Example: to measure $X \otimes X$, we can apply $\text{Had} \otimes \text{Had}$ to the quantum state, so that we can now measure in the computational basis ($Z \otimes Z$):

$$\langle \Psi | X \otimes X | \Psi \rangle = \langle \Psi | (\text{Had} \otimes \text{Had})(Z \otimes Z)(\text{Had} \otimes \text{Had}) | \Psi \rangle .$$

We will get a random output $\hat{m} \in \{0, 1\}$, such that

$$\mathbb{E}[-1^{\hat{m}}] = \langle \Psi | X \otimes X | \Psi \rangle .$$

Taking average over $N_s$ samples, the variance is $O(1/N_s)$. 
We need to do this for all terms. Can measure some of them simultaneously (e.g. for $X \otimes X$ and $Z \otimes Z$ because they commute), but this creates correlated error.
We need to do this for all terms. Can measure some of them simultaneously (e.g. for $X \otimes X$ and $Z \otimes Z$ because they commute), but this creates correlated error.

The total number of measurements to reach $\epsilon$ precision for $H$ is

$$\frac{(\text{some norm of } H)^2}{\epsilon^2}.$$ 

And we require roughly this many copies of $|\Psi\rangle$ (measurement destroys the quantum state).
We need to do this for all terms. Can measure some of them simultaneously (e.g. for $X \otimes X$ and $Z \otimes Z$ because they commute), but this creates correlated error.

The total number of measurements to reach $\epsilon$ precision for $H$ is

$$\frac{(\text{some norm of } H)^2}{\epsilon^2}.$$

And we require roughly this many copies of $|\Psi\rangle$ (measurement destroys the quantum state).

Quantum phase estimation can do the same by evolving with $H$ for $O(\epsilon^{-1})$ time (Heisenberg limit), with a single copy of $|\Psi\rangle$. 
Algorithms for different development stages of QC

- **NISQ**: variational algorithms (VQE, QAOA) ($\epsilon^{-2}$)
  - Need to optimize the total runtime, circuit depth, and number of qubits.

- **Fully fault-tolerant quantum algorithms** ($\epsilon^{-1}$)
  - Mainly consider the total runtime (dominated by non-Clifford gates), parallelization, energy consumption, etc.
Algorithms for different development stages of QC

- **NISQ**: variational algorithms (VQE, QAOA) ($\epsilon^{-2}$)

- **Early fault-tolerant quantum algorithms**
  - Need to optimize the total runtime, circuit depth, and number of qubits.

- **Fully fault-tolerant quantum algorithms**
  - Mainly consider the total runtime (dominated by non-Clifford gates), parallelization, energy consumption, etc.
Algorithms for different development stages of QC

- NISQ: variational algorithms (VQE, QAOA) ($\epsilon^{-2}$)

- Early fault-tolerant quantum algorithms
  - Need to optimize the total runtime, circuit depth, and number of qubits.

- Fully fault-tolerant quantum algorithms ($\epsilon^{-1}$)
  - Mainly consider the total runtime (dominated by non-Clifford gates), parallelization, energy consumption, etc.
Algorithms for different development stages of QC

- **NISQ**: variational algorithms (VQE, QAOA) \( (\epsilon^{-2}) \)

- **Fully fault-tolerant quantum algorithms** \( (\epsilon^{-1}) \)
  - Mainly consider the total runtime (dominated by non-Clifford gates), parallelization, energy consumption, etc.
Algorithms for different development stages of QC

- **NISQ**: variational algorithms (VQE, QAOA) ($\epsilon^{-2}$)

- Early fault-tolerant quantum algorithms
  - Need to optimize the total runtime, circuit depth, and number of qubits.

- **Fully fault-tolerant quantum algorithms** ($\epsilon^{-1}$)
  - Mainly consider the total runtime (dominated by non-Clifford gates), parallelization, energy consumption, etc.
The Heisenberg limit

▶ The quantum version of parameter estimation: estimate $\theta$ from parameterized quantum state $\rho(\theta)$, $\left\| \frac{d\rho}{d\theta} \right\|_1 \leq 1$ (here $\left\| \cdot \right\|_1$ is the trace norm).
The quantum version of parameter estimation: estimate $\theta$ from parameterized quantum state $\rho(\theta)$, $\| \frac{d\rho}{d\theta} \|_1 \leq 1$ (here $\| \cdot \|_1$ is the trace norm).

Information-theoretic lower bound: this requires $\Omega(\epsilon^{-2})$ samples (the standard quantum limit, SQL).
The Heisenberg limit

- The quantum version of **parameter estimation**: estimate $\theta$ from parameterized quantum state $\rho(\theta)$, $\|\frac{d\rho}{d\theta}\|_1 \leq 1$ (here $\|\cdot\|_1$ is the trace norm).

- Information-theoretic lower bound: this requires $\Omega(\epsilon^{-2})$ samples (**the standard quantum limit, SQL**).

- **Beyond-SQL example**: estimate eigenvalue to precision $\epsilon$ using QPE with exact eigenstate requires runtime $O(\epsilon^{-1})$. 
The quantum version of parameter estimation: estimate $\theta$ from parameterized quantum state $\rho(\theta)$, $\| \frac{d\rho}{d\theta} \|_1 \leq 1$ (here $\| \cdot \|_1$ is the trace norm).

Information-theoretic lower bound: this requires $\Omega(\epsilon^{-2})$ samples (the standard quantum limit, SQL).

Beyond-SQL example: estimate eigenvalue to precision $\epsilon$ using QPE with exact eigenstate requires runtime $O(\epsilon^{-1})$.

Information theoretic lower bound: this requires $\Omega(\epsilon^{-1})$ total evolution time (how long we evolve with $H$). This is the Heisenberg limit.
We have a target Hamiltonian $H$ (Hermitian matrix of size $2^N \times 2^N$, $N$ is the number of qubits).

$\lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots$, $\lambda_0$ is the ground state energy. $|\Psi_0\rangle$ is the ground state. $|\Phi\rangle$ is an initial guess for the ground state. We can apply control $-i\tau H$, where $\tau$ is a rescaling factor.
We have a target Hamiltonian $H$ (Hermitian matrix of size $2^N \times 2^N$, $N$ is the number of qubits).

- Its eigenvalues and corresponding eigenstates are $\lambda_k$ and $|\Psi_k\rangle$ respectively.

- $|\Phi\rangle$ is an initial guess for the ground state.

- We can apply control-$e^{-i\tau H}$, where $\tau$ is a rescaling factor.
We have a target Hamiltonian $H$ (Hermitian matrix of size $2^N \times 2^N$, $N$ is the number of qubits).

Its eigenvalues and corresponding eigenstates are $\lambda_k$ and $|\Psi_k\rangle$ respectively.

$\lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$, $\lambda_0$ is the ground state energy. $|\Psi_0\rangle$ is the ground state.
We have a target Hamiltonian $H$ (Hermitian matrix of size $2^N \times 2^N$, $N$ is the number of qubits).

Its eigenvalues and corresponding eigenstates are $\lambda_k$ and $|\Psi_k\rangle$ respectively.

$\lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$, $\lambda_0$ is the ground state energy. $|\Psi_0\rangle$ is the ground state.

$|\Phi\rangle$ is an initial guess for the ground state.
We have a target Hamiltonian $H$ (Hermitian matrix of size $2^N \times 2^N$, $N$ is the number of qubits).

- Its eigenvalues and corresponding eigenstates are $\lambda_k$ and $|\Psi_k\rangle$ respectively.
- $\lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$, $\lambda_0$ is the ground state energy. $|\Psi_0\rangle$ is the ground state.
- $|\Phi\rangle$ is an initial guess for the ground state.
- We can apply control-$e^{-i\tau H}$, where $\tau$ is a rescaling factor.
We will use the following asymptotic notations:

- $f(x) = \mathcal{O}(g(x))$ if there exists $C > 0$ such that $f(x) \leq C g(x)$ (for $x$ larger/smaller than some threshold).

- $f(x) = e\mathcal{O}(g(x))$ if $f(x) = \mathcal{O}(g(x) \text{ polylog}(g(x)))$.

- $f(x) = \Omega(g(x))$ if $g(x) = \mathcal{O}(f(x))$.

- $f(x) = e\Omega(g(x))$ if $f(x) = \Omega(g(x)/\text{polylog}(g(x)))$. 
We will use the following asymptotic notations:

- $f(x) = \mathcal{O}(g(x))$ if there exists $C > 0$ such that $f(x) \leq Cg(x)$ (for $x$ larger/smaller than some threshold).

- $f(x) = \tilde{\mathcal{O}}(g(x))$ if $f(x) = \mathcal{O}(g(x) \text{polylog}(g(x)))$. 

We will use the following asymptotic notations:

- $f(x) = \mathcal{O}(g(x))$ if there exists $C > 0$ such that $f(x) \leq C g(x)$ (for $x$ larger/smaller than some threshold).

- $f(x) = \widetilde{\mathcal{O}}(g(x))$ if $f(x) = \mathcal{O}(g(x) \text{polylog}(g(x)))$.

- $f(x) = \Omega(g(x))$ if $g(x) = \mathcal{O}(f(x))$. 
We will use the following asymptotic notations:

- \( f(x) = \mathcal{O}(g(x)) \) if there exists \( C > 0 \) such that \( f(x) \leq C g(x) \) (for \( x \) larger/smaller than some threshold).

- \( f(x) = \tilde{\mathcal{O}}(g(x)) \) if \( f(x) = \mathcal{O}(g(x) \text{ polylog}(g(x))) \).

- \( f(x) = \Omega(g(x)) \) if \( g(x) = \mathcal{O}(f(x)) \).

- \( f(x) = \tilde{\Omega}(g(x)) \) if \( f(x) = \Omega(g(x) / \text{ polylog}(g(x))) \).
Single-qubit quantum phase estimation

Figure: The Hadamard test circuit: from the measurement outcome $\hat{m}$ we can compute the expectation value $\langle \Phi | e^{-itH} | \Phi \rangle$. Real and imaginary parts are computed separately (corresponding to $\text{Had}$ and $\text{Had}S^\dagger$ respectively).
For the real part, before measurement the quantum state undergoes the transformation

\[ |0\rangle |\Phi\rangle \rightarrow |+\rangle |\Phi\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle |\Phi\rangle + |1\rangle e^{-itH} |\Phi\rangle) \]
For the real part, before measurement the quantum state undergoes the transformation

\[ |0\rangle |\Phi\rangle \rightarrow |+\rangle |\Phi\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle |\Phi\rangle + |1\rangle e^{-itH} |\Phi\rangle) \]

\[ \rightarrow \frac{1}{\sqrt{2}}(|+\rangle |\Phi\rangle + |\rangle e^{-itH} |\Phi\rangle) \]
For the real part, before measurement the quantum state undergoes the transformation

\[
|0\rangle |\Phi\rangle \mapsto |+\rangle |\Phi\rangle \mapsto \frac{1}{\sqrt{2}}(|0\rangle |\Phi\rangle + |1\rangle e^{-itH} |\Phi\rangle)
\]

\[
\mapsto \frac{1}{\sqrt{2}}(|+\rangle |\Phi\rangle + |\rangle e^{-itH} |\Phi\rangle)
\]

\[
= \frac{1}{2}((|0\rangle + |1\rangle) |\Phi\rangle + (|0\rangle - |1\rangle) e^{-itH} |\Phi\rangle)
\]
For the real part, before measurement the quantum state undergoes the transformation

\[|0\rangle|\Phi\rangle \rightarrow |+\rangle|\Phi\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle|\Phi\rangle + |1\rangle e^{-itH}|\Phi\rangle)\]

\[\rightarrow \frac{1}{\sqrt{2}}(|+\rangle|\Phi\rangle + |-\rangle e^{-itH}|\Phi\rangle)\]

\[= \frac{1}{2}((|0\rangle + |1\rangle)|\Phi\rangle + (|0\rangle - |1\rangle)e^{-itH}|\Phi\rangle)\]

\[= \frac{1}{2}(|0\rangle(|\Phi\rangle + e^{-itH}|\Phi\rangle) + |1\rangle(|\Phi\rangle - e^{-itH}|\Phi\rangle))\]
For the real part, before measurement the quantum state undergoes the transformation

\[
|0\rangle |\Phi\rangle \rightarrow |+\rangle |\Phi\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle |\Phi\rangle + |1\rangle e^{-itH} |\Phi\rangle)
\]

\[
\rightarrow \frac{1}{\sqrt{2}}(|+\rangle |\Phi\rangle + |-\rangle e^{-itH} |\Phi\rangle)
\]

\[
= \frac{1}{2}((|0\rangle + |1\rangle) |\Phi\rangle + (|0\rangle - |1\rangle) e^{-itH} |\Phi\rangle)
\]

\[
= \frac{1}{2}(|0\rangle (|\Phi\rangle + e^{-itH} |\Phi\rangle) + |1\rangle (|\Phi\rangle - e^{-itH} |\Phi\rangle))
\]

We then measure the first qubit to get \(\hat{m} \in \{0, 1\}\). The expectation value of \((-1)^{\hat{m}}\) is

\[
\mathbb{E}[(-1)^{\hat{m}}] = \frac{1}{4}(\| |\Phi\rangle + e^{-itH} |\Phi\rangle \|^2 - \| |\Phi\rangle - e^{-itH} |\Phi\rangle \|^2) = \text{Re} \langle \Phi | e^{-itH} |\Phi\rangle.
\]
Similarly we can get the imaginary part.

For any $t$, we estimate $\langle \Phi | e^{-itH} | \Phi \rangle$ by

$$S(t) = \langle \Phi | e^{-itH} | \Phi \rangle + e(t),$$

where $e(t)$ is statistical noise.

This signal contains eigenvalue information

$$\langle \Phi | e^{-itH} | \Phi \rangle = \sum_k X_k e^{-it\lambda_k} |\langle \Phi | \Psi_k \rangle|^2.$$ 

Obtain eigenvalues of $H$ by processing the signal and getting the target frequency.
Similarly we can get the imaginary part.

For any $t$, we estimate $\langle \Phi | e^{-itH} | \Phi \rangle$ by

$$S(t) = \langle \Phi | e^{-itH} | \Phi \rangle + e(t),$$

where $e(t)$ is statistical noise.
Similarly we can get the imaginary part.

For any $t$, we estimate $\langle \Phi | e^{-itH} | \Phi \rangle$ by

$$S(t) = \langle \Phi | e^{-itH} | \Phi \rangle + e(t),$$

where $e(t)$ is statistical noise.

This signal contains eigenvalue information

$$\langle \Phi | e^{-itH} | \Phi \rangle = \sum_k e^{-it\lambda_k} | \langle \Phi | \Psi_k \rangle |^2.$$
Similarly we can get the imaginary part.

For any $t$, we estimate $\langle \Phi | e^{-itH} | \Phi \rangle$ by

$$S(t) = \langle \Phi | e^{-itH} | \Phi \rangle + e(t),$$

where $e(t)$ is statistical noise.

This signal contains eigenvalue information

$$\langle \Phi | e^{-itH} | \Phi \rangle = \sum_k e^{-it\lambda_k} | \langle \Phi | \Psi_k \rangle |^2.$$

Obtain eigenvalues of $H$ by processing the signal and getting the target frequency.
Single frequency estimation

We have seen that the Hadamard test circuit can be used to compute

\[ \langle \Phi | e^{-itH} | \Phi \rangle = \sum_k e^{-it\lambda_k} \langle \Phi | \Psi_k \rangle^2. \]
We have seen that the Hadamard test circuit can be used to compute

\[ \langle \Phi | e^{-itH} | \Phi \rangle = \sum_k e^{-it\lambda_k} | \langle \Phi | \Psi_k \rangle |^2. \]

The simplest case: if \( |\Phi\rangle = |\Psi_0\rangle \), then we can use this to get the ground state energy \( \lambda_0 \).

\[ S(t) = e^{-i\lambda_0 t} + e(t). \]
From the Hadamard test circuit we can generate $S(t)$, $t \geq 0$

$$S(t) = e^{-i\lambda_0 t} + e(t).$$

We want to estimate $\lambda_0 \in (-1, 1]$ (rescaling the Hamiltonian properly) to precision $\epsilon$.

\[^2\text{Kimmel, Low, Yoder, 2015, Robust Calibration of a Universal Single-Qubit Gate-Set via Robust Phase Estimation.}\]
From the Hadamard test circuit we can generate $S(t)$, $t \geq 0$

$$S(t) = e^{-i\lambda_0 t} + e(t).$$

We want to estimate $\lambda_0 \in (-1, 1]$ (rescaling the Hamiltonian properly) to precision $\epsilon$.

- We can let $t = \pi/2$, average out the noise, and estimate $\theta$ with $O(\epsilon^{-2})$ samples.
From the Hadamard test circuit we can generate $S(t)$, $t \geq 0$

$$S(t) = e^{-i\lambda_0 t} + e(t).$$

We want to estimate $\lambda_0 \in (-1, 1]$ (rescaling the Hamiltonian properly) to precision $\epsilon$.

- We can let $t = \pi/2$, average out the noise, and estimate $\theta$ with $\mathcal{O}(\epsilon^{-2})$ samples.

- I will outline a method that uses (ignoring the log log factor)$^2$
  
  1. $\mathcal{O}(\log(\epsilon^{-1}))$ samples,
  2. $\mathcal{O}(\epsilon^{-1})$ total evolution time.

---

$^2$Kimmel, Low, Yoder, 2015, *Robust Calibration of a Universal Single-Qubit Gate-Set via Robust Phase Estimation*. 
From the Hadamard test circuit we can generate $S(t)$, $t \geq 0$

$$S(t) = e^{-i \lambda_0 t} + e(t).$$

We want to estimate $\lambda_0 \in (-1, 1]$ (rescaling the Hamiltonian properly) to precision $\epsilon$.

- We can let $t = \pi/2$, average out the noise, and estimate $\theta$ with $O(\epsilon^{-2})$ samples.

- I will outline a method that uses (ignoring the $\log \log$ factor)$^2$
  
  1. $O(\log(\epsilon^{-1}))$ samples,
  2. $O(\epsilon^{-1})$ total evolution time.

- Suppose our samples are $S(t_1), S(t_2), \cdots, S(t_{N_s})$, then the total evolution time is $t_1 + t_2 + \cdots + t_{N_s}$.

---

$^2$Kimmel, Low, Yoder, 2015, *Robust Calibration of a Universal Single-Qubit Gate-Set via Robust Phase Estimation.*
Suppose we know $a \leq -\lambda_0 \leq b$. We want to determine

1. $a \leq -\lambda_0 \leq \frac{a+2b}{3},$
2. or $\frac{2a+b}{3} \leq -\lambda_0 \leq b.$
Suppose we know $a \leq -\lambda_0 \leq b$. We want to determine

1. $a \leq -\lambda_0 \leq \frac{a+2b}{3}$,
2. or $\frac{2a+b}{3} \leq -\lambda_0 \leq b$.

If we can do that then we can reduce the uncertainty by $1/3$ at each step. $O(\log(\epsilon^{-1}))$ steps are needed for $\epsilon$ precision.
We look at the value of

\[ f_{a,b}(-\lambda_0) = \sin \left( \frac{\pi}{b-a} \left( -\lambda_0 - \frac{a+b}{2} \right) \right) = \text{Im} \langle S(t^*) \rangle e^{-i \frac{(a+b)\pi}{2(b-a)}}, \]

where \( t^* = \frac{\pi}{b-a} \).
We look at the value of

\[ f_{a,b}(-\lambda_0) = \sin \left( \frac{\pi}{b-a} \left( -\lambda_0 - \frac{a + b}{2} \right) \right) = \text{Im} \langle S(t^*) \rangle e^{-i \frac{(a+b)\pi}{2(b-a)}} , \]

where \( t^* = \frac{\pi}{b-a} \).
We look at the value of

\[ f_{a,b}(-\lambda_0) = \sin \left( \frac{\pi}{b-a} \left( -\lambda_0 - \frac{a+b}{2} \right) \right) = \text{Im} \langle S(t^*) \rangle e^{-i \frac{(a+b)\pi}{2(b-a)}} , \]

where \( t^* = \frac{\pi}{b-a} \).

\[ \text{If } f_{a,b}(-\lambda_0) \leq \frac{1}{2}, \text{ then } a \leq -\lambda_0 \leq \frac{a+2b}{3} ; \]
We look at the value of
\[ f_{a,b}(-\lambda_0) = \sin\left( \frac{\pi}{b-a} \left( -\lambda_0 - \frac{a+b}{2} \right) \right) = \text{Im} \left\langle S(t^*) \right\rangle e^{-i \frac{(a+b)\pi}{2(b-a)}}, \]
where \( t^* = \frac{\pi}{b-a} \).

\[ \begin{align*}
\text{▷ If } f_{a,b}(-\lambda_0) &\leq \frac{1}{2}, \text{ then } a \leq -\lambda_0 \leq \frac{a+2b}{3}; \\
\text{▷ If } f_{a,b}(-\lambda_0) &\geq -\frac{1}{2}, \text{ then } \frac{2a+b}{3} \leq -\lambda_0 \leq b. 
\end{align*} \]

\[ a \quad (2a+b)/3 \quad (a+2b)/3 \quad b \]
We look at the value of

\[ f_{a,b}(-\lambda_0) = \sin \left( \frac{\pi}{b-a} \left( -\lambda_0 - \frac{a+b}{2} \right) \right) = \text{Im} \langle S(t^*) \rangle e^{-i \frac{(a+b)\pi}{2(b-a)}}, \]

where \( t^* = \frac{\pi}{b-a} \).

- If \( f_{a,b}(-\lambda_0) \leq \frac{1}{2} \), then \( a \leq -\lambda_0 \leq \frac{a+2b}{3} \);
- If \( f_{a,b}(-\lambda_0) \geq -\frac{1}{2} \), then \( \frac{2a+b}{3} \leq -\lambda_0 \leq b \);
- Evaluating \( f_{a,b}(-\lambda_0) \) to precision \( \frac{1}{2} \) is enough.
We look at the value of

\[ f_{a,b}(-\lambda_0) = \sin \left( \frac{\pi}{b-a} \left( -\lambda_0 - \frac{a+b}{2} \right) \right) = \text{Im} \langle S(t^*) \rangle e^{-i \frac{(a+b)\pi}{2(b-a)}} , \]

where \( t^* = \frac{\pi}{b-a} \).

\[ \text{▶ If } f_{a,b}(-\lambda_0) \leq \frac{1}{2}, \text{ then } a \leq -\lambda_0 \leq \frac{a+2b}{3} ; \]

\[ \text{▶ If } f_{a,b}(-\lambda_0) \geq -\frac{1}{2}, \text{ then } \frac{2a+b}{3} \leq -\lambda_0 \leq b. \]

\[ \text{▶ Evaluating } f_{a,b}(-\lambda_0) \text{ to precision } \frac{1}{2} \text{ is enough.} \]

\[ \text{▶ Can get confidence level } 1 - \delta' \text{ with } \mathcal{O}(\log(\delta'^{-1})) \text{ samples.} \]
At the last search step $b - a \approx (3/2)\epsilon$, and therefore $t^* \approx (2/3)\epsilon^{-1}$. 
At the last search step \( b - a \approx (3/2)\epsilon \), and therefore \( t^* \approx (2/3)\epsilon^{-1} \).

The cost of the last step is \( \mathcal{O}(t^* \log(\delta'^{-1})) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})) \).
At the last search step $b - a \approx (3/2)\epsilon$, and therefore $t^* \approx (2/3)\epsilon^{-1}$.

The cost of the last step is $\mathcal{O}(t^* \log(\delta'^{-1})) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1}))$.

The total cost is

$$\mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})) \times \left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \cdots\right) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})).$$

This achieves the Heisenberg-limited scaling.
At the last search step $b - a \approx (3/2)\epsilon$, and therefore $t^* \approx (2/3)\epsilon^{-1}$.

The cost of the last step is $\mathcal{O}(t^* \log(\delta'^{-1})) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1}))$.

The total cost is

$$\mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})) \times \left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \cdots \right) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})).$$

This achieves the Heisenberg-limited scaling.

Need $\delta' = \mathcal{O}(\delta/\log(\epsilon^{-1}))$ to ensure that all steps are successful with probability $1 - \delta$. 
At the last search step \( b - a \approx (3/2)\epsilon \), and therefore \( t^* \approx (2/3)\epsilon^{-1} \).

The cost of the last step is \( \mathcal{O}(t^* \log(\delta'^{-1})) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})) \).

The total cost is

\[
\mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})) \times \left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \cdots\right) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})).
\]

This achieves the Heisenberg-limited scaling.

Need \( \delta' = \mathcal{O}(\delta / \log(\epsilon^{-1})) \) to ensure that all steps are successful with probability \( 1 - \delta \).

Total evolution time is \( \mathcal{O}(\epsilon^{-1} \log(\delta^{-1})) \) and the number of samples is \( \mathcal{O}(\log(\epsilon^{-1})) \).

Robust to noise (\( |e(t)| \leq 1/2 \) w.p. 2/3).
A high-level statement: We can use the Hadamard test circuit to estimate the eigenvalue given the corresponding eigenstate with Heisenberg-limited scaling. It is also robust to constant amount of noise.
A high-level statement: We can use the Hadamard test circuit to estimate the eigenvalue given the corresponding eigenstate with Heisenberg-limited scaling. It is also robust to constant amount of noise.

A precise statement: we can do the above with $O(\epsilon^{-1} \log(\delta^{-1}))$ total evolution time to get confidence level $1 - \delta$. We will still get correct estimate when $|e(t)| \leq 1/2$ w.p. $2/3$. 
Quantum phase estimation

\[ |0\rangle \xrightarrow{\text{Had}} \cdots \xrightarrow{\text{QFT}\dagger} \]

\[ |0\rangle \xrightarrow{\text{Had}} \cdots \xrightarrow{\text{QFT}\dagger} \]

\[ |0\rangle \xrightarrow{\text{Had}} \cdots \xrightarrow{\text{QFT}\dagger} \]

\[ |\Phi\rangle \xrightarrow{e^{-i\tau H}} e^{-2i\tau H} \cdots e^{-2^{r-1}i\tau H} \]

**Figure:** The quantum phase estimation circuit. Two registers: energy register (r qubits) and state register (N qubits). Measuring the energy register yields a bit string \( \hat{m} \), which we convert to an energy measurement \( \tau \hat{\lambda} = 2\pi \hat{m} / 2^r \).
Initial state is $|00\cdots0\rangle|\Phi\rangle$. 
Initial state is $|00\cdots0\rangle|\Phi\rangle$.

Apply Hadamard gates:

$$|++\cdots\rangle|\Phi\rangle = \frac{1}{\sqrt{2^r}} \sum_{j=0}^{2^r-1} |j\rangle|\Phi\rangle.$$
Initial state is $|00\cdots 0\rangle |\Phi\rangle$.

Apply Hadamard gates:

$$|++\cdots +\rangle |\Phi\rangle = \frac{1}{\sqrt{2^r}} \sum_{j=0}^{2^r-1} |j\rangle |\Phi\rangle.$$ 

Controlled time evolution:

$$\frac{1}{\sqrt{2^r}} \sum_{j=0}^{2^r-1} |j\rangle |\Phi\rangle \mapsto \frac{1}{\sqrt{2^r}} \sum_{j=0}^{2^r-1} |j\rangle e^{-ij\tau H} |\Phi\rangle.$$
Eigenbasis expansion: let $|\Phi\rangle = \sum_k c_k |\Psi_k\rangle$

$$\frac{1}{\sqrt{2^r}} \sum_{j=0}^{r-1} |j\rangle e^{-ij \tau H} |\Phi\rangle = \sum_k c_k \frac{1}{\sqrt{2^r}} \sum_{j=0}^{r-1} |j\rangle e^{-ij \tau \lambda_k} |\Psi_k\rangle.$$
Eigenbasis expansion: let $|\Phi\rangle = \sum_k c_k |\Psi_k\rangle$

$$\frac{1}{\sqrt{2^r}} \sum_{j=0}^{r-1} |j\rangle e^{-ij\tau H} |\Phi\rangle = \sum_k c_k \frac{1}{\sqrt{2^r}} \sum_{j=0}^{r-1} |j\rangle e^{-ij\tau \lambda_k} |\Psi_k\rangle.$$ 

Inverse QFT: $|j\rangle \mapsto \frac{1}{\sqrt{2^r}} \sum_{m=0}^{2^r-1} e^{i2\pi jm/2^r} |m\rangle.$

$$\sum_k c_k \sum_{m=0}^{2^r-1} |m\rangle \frac{1}{2^r} \sum_{j=0}^{r-1} e^{i2\pi jm/2^r - ij\tau \lambda_k} |\Psi_k\rangle \Gamma(2\pi m/2^r - \tau \lambda_k)$$
The reason why QPE works: $\Gamma(\theta)$ approximates the Dirac delta function.

$$\sum_k c_k \sum_{m=0}^{2^r-1} |m\rangle \Gamma(2\pi m/2^r - \tau \lambda_k) |\Psi_k\rangle$$

$$\approx \sum_k c_k \sum_{m=0}^{2^r-1} |m\rangle \delta(2\pi m/2^r - \tau \lambda_k) |\Psi_k\rangle$$

$$= \sum_k c_k \left| \frac{2^r \tau \lambda_k}{2\pi} \right\rangle |\Psi_k\rangle$$

From the energy register we can directly read off $2^r \tau \lambda_k$ (binary representation of rescaled $\lambda_k$).

This is the idealized version of QPE. Next we will see what actually happens.
The reason why QPE works: \( \Gamma(\theta) \) approximates the Dirac delta function.

\[
\sum_k c_k \sum_{m=0}^{2^r-1} |m\rangle \Gamma(2\pi m/2^r - \tau \lambda_k) |\Psi_k\rangle 
\]

\[
\approx \sum_k c_k \sum_{m=0}^{2^r-1} |m\rangle \delta(2\pi m/2^r - \tau \lambda_k) |\Psi_k\rangle 
\]

\[
= \sum_k c_k |\frac{2^r \tau \lambda_k}{2\pi}\rangle |\Psi_k\rangle 
\]

From the energy register we can directly read off \( \frac{2^r \tau \lambda_k}{2\pi} \) (binary representation of rescaled \( \lambda_k \)).
The reason why QPE works: $\Gamma(\theta)$ approximates the Dirac delta function.

$$\sum_k c_k \sum_{m=0}^{2^r-1} |m\rangle \Gamma\left(2\pi m/2^r - \tau \lambda_k\right) |\Psi_k\rangle$$

$$\approx \sum_k c_k \sum_{m=0}^{2^r-1} |m\rangle \delta\left(2\pi m/2^r - \tau \lambda_k\right) |\Psi_k\rangle$$

$$= \sum_k c_k \left|\frac{2^r \tau \lambda_k}{2\pi}\right| |\Psi_k\rangle$$

From the energy register we can directly read off $\frac{2^r \tau \lambda_k}{2\pi}$ (binary representation of rescaled $\lambda_k$).

This is the idealized version of QPE. Next we will see what actually happens.
The kernel function (resembling the Dirac delta function)

\[ \Gamma(\theta) = \frac{1}{2^r} \sum_{j=0}^{r-1} e^{ij\theta} = \frac{1}{2^r} \frac{1 - e^{i2^r\theta}}{1 - e^{i\theta}}. \]
The kernel function (resembling the Dirac delta function)

\[ \Gamma(\theta) = \frac{1}{2^r} \sum_{j=0}^{r-1} e^{ij\theta} = \frac{1}{2^r} \frac{1 - e^{i2^r\theta}}{1 - e^{i\theta}}. \]

Simplify the quantum state

\[ \sum_{k} c_k \sum_{m=0}^{2^r-1} |m\rangle \Gamma(\frac{2\pi m}{2^r} - \tau \lambda_k) |\Psi_k\rangle. \]
The kernel function (resembling the Dirac delta function)

\[ \Gamma(\theta) = \frac{1}{2r} \sum_{j=0}^{r-1} e^{ij\theta} = \frac{1}{2r} \frac{1 - e^{i2r\theta}}{1 - e^{i\theta}}. \]

Simplify the quantum state

\[ \sum_{k} c_k \sum_{m=0}^{2^r-1} |m\rangle \Gamma(2\pi m / 2^r - \tau \lambda_k) |\Psi_k\rangle. \]

Measuring the energy register \( |m\rangle \) yields \( m \) with probability

\[ \Pr[\hat{m} = m] = \sum_{k} |c_k|^2 |\Gamma(2\pi m / 2^r - \tau \lambda_k)|^2. \]
Suppose we are given a hidden random variable $\hat{k}$, with

$$\Pr[\hat{k} = k] = |c_k|^2.$$  

Then

$$\Pr[\hat{m} = m] = \sum_k \Pr[\hat{k} = k] \Pr[\hat{m} = m | \hat{k} = k],$$

where

$$\Pr[\hat{m} = m | \hat{k} = k] = |\Gamma(2\pi m/2^r - \tau \lambda_k)|^2.$$
Suppose we are given a hidden random variable $\hat{k}$, with

$$\Pr[\hat{k} = k] = |c_k|^2.$$  

Then

$$\Pr[\hat{m} = m] = \sum_k \Pr[\hat{k} = k] \Pr[\hat{m} = m|\hat{k} = k],$$

where

$$\Pr[\hat{m} = m|\hat{k} = k] = |\Gamma(2\pi m/2^r - \tau \lambda_k)|^2.$$  

$k$ is the index for the eigenstate. $\Pr[\hat{m} = m|\hat{k} = k]$ is the probability of getting energy measurement $\hat{m}$ given an eigenstate $|\Psi_k\rangle$. 
Suppose we are given a \textit{hidden} random variable $\hat{k}$, with

$$\Pr[\hat{k} = k] = |c_k|^2.$$  

Then

$$\Pr[\hat{m} = m] = \sum_k \Pr[\hat{k} = k] \Pr[\hat{m} = m|\hat{k} = k],$$

where

$$\Pr[\hat{m} = m|\hat{k} = k] = |\Gamma(2\pi m/2^r - \tau \lambda_k)|^2.$$  

$k$ is the index for the eigenstate. $\Pr[\hat{m} = m|\hat{k} = k]$ is the probability of getting energy measurement $\hat{m}$ given an eigenstate $|\Psi_k\rangle$.

Need to show that $\Pr[\hat{m} = m|\hat{k} = k]$ is concentrated around $\frac{2^r \tau \lambda_k}{2\pi}$. 
We let

\[ \Delta \theta = \frac{2\pi m}{2r} - \tau \lambda_k, \]

\begin{align*}
\text{energy measurement} & \quad \text{rescaled energy}
\end{align*}
We let

\[ \Delta \theta = \frac{2\pi m}{2^r} - \tau \lambda_k, \]

energy measurement  rescaled energy

then

\[
\Pr[\hat{m} = m | \hat{k} = k] = |\Gamma(2\pi m/2^r - \tau \lambda_k)|^2
\]

\[
= \left| \frac{1}{2^r} \frac{1 - e^{i2^r \Delta \theta}}{1 - e^{i\Delta \theta}} \right|^2
\]

\[
= \frac{1}{4^r} \sin^2(2^{r-1} \Delta \theta)
\]

\[
= \frac{1}{4^r} \sin^2(\Delta \theta/2).
\]
We let

$$\Delta \theta = \frac{2\pi m}{2^r} - \tau \lambda_k,$$

where energy measurement $\text{energy measurement}$ and rescaled energy $\text{rescaled energy}$, then

$$\Pr[\hat{m} = m | \hat{k} = k] = |\Gamma(2\pi m/2^r - \tau \lambda_k)|^2$$

$$= \left| \frac{1}{2^r} \frac{1 - e^{i2^r \Delta \theta}}{1 - e^{i\Delta \theta}} \right|^2$$

$$= \frac{1}{4^r} \frac{\sin^2(2^r-1 \Delta \theta)}{\sin^2(\Delta \theta/2)}.$$

**Figure:** $|\Gamma(\Delta \theta)|^2$ for $\Delta \theta \in [-\pi, \pi]$. The kernel is $2\pi$-periodic.
Define the energy measurement \( \tau \hat{\lambda} = 2\pi \hat{m}/2^r \) which we get from running the QPE circuit.
Define the energy measurement $\tau\hat{\lambda} = 2\pi\hat{m}/2^r$ which we get from running the QPE circuit.

Define distance on torus $|x|_a = \min_{k \in \mathbb{Z}} |x - ka|$. 
Define the energy measurement $\tau \hat{\lambda} = 2\pi \hat{m}/2^r$ which we get from running the QPE circuit.

Define distance on torus $|x|_a = \min_{k \in \mathbb{Z}} |x - ka|.$

The concentration of the kernel $\Gamma(\Delta \theta)$ guarantees that

$$\Pr[|\tau \hat{\lambda} - \tau \lambda_k|2\pi \geq \epsilon |\hat{k} = k] \leq O \left( \frac{1}{2^r \epsilon} \right).$$
An operational understanding of QPE

- We first sample $\hat{k} = k$ w.p. $|c_k|^2$ (remember $|\Phi\rangle = \sum_k c_k |\Psi_k\rangle$).
An operational understanding of QPE

- We first sample $\hat{k} = k$ w.p. $|c_k|^2$ (remember $|\Phi\rangle = \sum_k c_k |\Psi_k\rangle$).

- We do not actually do this and we have no access to $k$. 

- An energy estimate $\hat{\tau} = \lambda$ is generated from running QPE that is $\epsilon$-close to $\tau$ with probability at least $1 - O(\epsilon^{-1/2} - r)$.

- In this sense we are sampling from the spectrum of $\tau H$, and each sample is close to an (rescaled) eigenvalue with large probability (exact in the limit of $r \to \infty$).

- The rescaled eigenvalues are determined modulo $2\pi$. We need to choose $\tau$ appropriately to get $\lambda_k$. 

We first sample $\hat{k} = k$ w.p. $|c_k|^2$ (remember $|\Phi\rangle = \sum_k c_k |\Psi_k\rangle$).

We do not actually do this and we have no access to $k$.

An energy estimate $\tau \hat{\lambda}$ is generated from running QPE that is $\epsilon$-close to $\tau \lambda_k$ with probability at least $1 - O(\epsilon^{-1} 2^{-r})$. In this sense we are sampling from the spectrum of $\tau H$, and each sample is close to an (rescaled) eigenvalue with large probability (exact in the limit of $r \to \infty$).

The rescaled eigenvalues are determined modulo $2\pi$. We need to choose $\tau$ appropriately to get $\lambda_k$. 

An operational understanding of QPE
An operational understanding of QPE

- We first sample $\hat{k} = k$ w.p. $|c_k|^2$ (remember $|\Phi\rangle = \sum_k c_k |\Psi_k\rangle$).

- We do not actually do this and we have no access to $k$.

- An energy estimate $\tau\hat{\lambda}$ is generated from running QPE that is $\epsilon$-close to $\tau\lambda_k$ with probability at least $1 - \mathcal{O}(\epsilon^{-1}2^{-r})$.

- In this sense we are sampling from the spectrum of $\tau H$, and each sample is close to an (rescaled) eigenvalue with large probability (exact in the limit of $r \to \infty$).
An operational understanding of QPE

- We first sample $\hat{k} = k$ w.p. $|c_k|^2$ (remember $|\Phi\rangle = \sum_k c_k |\Psi_k\rangle$).

- We do not actually do this and we have no access to $k$.

- An energy estimate $\tau \hat{\lambda}$ is generated from running QPE that is $\epsilon$-close to $\tau \lambda_k$ with probability at least $1 - O(\epsilon^{-1} 2^{-r})$.

- In this sense we are sampling from the spectrum of $\tau H$, and each sample is close to an (rescaled) eigenvalue with large probability (exact in the limit of $r \to \infty$).

- The rescaled eigenvalues are determined modulo $2\pi$. We need to choose $\tau$ appropriately to get $\lambda_k$. 
Proof of the concentration inequality

- Want to prove $\text{Pr}[|\tau \hat{\lambda} - \tau \lambda_k|_{2\pi} \geq \epsilon | \hat{k} = k] \leq O\left(\frac{1}{2^r \epsilon}\right)$. 
Proof of the concentration inequality

▶ Want to prove \( \Pr[|\tau \hat{\lambda} - \tau \lambda_k|_{2\pi} \geq \epsilon | \hat{k} = k] \leq O \left( \frac{1}{2^r \epsilon} \right) \).

▶ The probability of deviation larger than \( \epsilon = 2\pi \ell/2^r \)

\[
\Pr \left[ \left| \frac{2\pi \hat{m}}{2^r} - \tau \lambda_k \right|_{2\pi} \geq \frac{2\pi \ell}{2^r} | \hat{k} = k \right]
= \sum_{m: m \sim \tau \lambda_k \frac{2^r}{2\pi}} \frac{1}{4^r} \frac{\sin^2 \left( \frac{2^r - 1}{2} (2\pi m/2^r - \tau \lambda_k) \right)}{\sin^2 \left( \frac{(2\pi m/2^r - \tau \lambda_k)}{2} \right)}
\]
Proof of the concentration inequality

- Want to prove \( \Pr[|\tau \hat{\lambda} - \tau \lambda_k|_{2\pi} \geq \epsilon | \hat{k} = k] \leq O \left( \frac{1}{2^r \epsilon} \right) \).

- The probability of deviation larger than \( \epsilon = 2\pi \ell / 2^r \)

\[
\Pr \left[ \left| \frac{2\pi \hat{m}}{2^r} - \tau \lambda_k \right|_{2\pi} \geq \frac{2\pi \ell}{2^r} \right| \hat{k} = k \] = \sum_{m: \left| m - \frac{2^r \tau \lambda_k}{2\pi} \right|_{2^r} \geq \ell} \left[ \frac{1}{4^r} \frac{1}{\sin^2 \left( \frac{(2\pi m/2^r - \tau \lambda_k)}/2 \right)} \right] 
\leq \sum_{m: \left| m - \frac{2^r \tau \lambda_k}{2\pi} \right|_{2^r} \geq \ell} \left[ \frac{1}{4^r} \frac{1}{\sin^2 \left( \frac{(2\pi m/2^r - \tau \lambda_k)}/2 \right)} \right]
We use the fact that $|\sin \left( \frac{x}{2} \right)| \geq \frac{|x|_{2\pi}}{\pi}$. 
We use the fact that $|\sin \left( \frac{x}{2} \right)| \geq \frac{|x|}{2\pi} \cdot$

$$\sum_{m: \left|m - \frac{2^r \tau\lambda_k}{2\pi}\right| \geq \ell} \frac{1}{4^r \sin^2 \left( (2\pi m/2^r - \tau\lambda_k)/2 \right)}$$
We use the fact that \( |\sin \left( \frac{x}{2} \right) | \geq \frac{|x|}{2\pi} \).

\[
\sum_{m : \left| m - \frac{2^r \tau \lambda_k}{2\pi} \right|_{2^r} \geq \ell} \frac{1}{4^r \sin^2 \left( (\frac{2\pi m}{2^r} - \tau \lambda_k) / 2 \right)} \leq \sum_{m : \left| m - \frac{2^r \tau \lambda_k}{2\pi} \right|_{2^r} \geq \ell} \frac{1}{4|m - 2^{r-1} \tau \lambda_k / \pi|^2_{2^r}}
\]
We use the fact that $|\sin \left( \frac{x}{2} \right)| \geq \frac{|x|_{2\pi}}{\pi}$.

\[
\sum_{m: \left| m - \frac{2^r \tau \lambda_k}{2\pi} \right|_{2^r \pi} \geq \ell} \frac{1}{4r} \frac{1}{\sin^2 \left( \frac{(2\pi m / 2^r - \tau \lambda_k)/2}{2} \right)}
\leq \sum_{m: \left| m - \frac{2^r \tau \lambda_k}{2^r \pi} \right|_{2\pi} \geq \ell} \frac{1}{4\left| m - 2^r - 1 \tau \lambda_k \pi \right|_{2^r}}
\leq 2 \times \frac{1}{4} \sum_{n=\ell}^{\infty} \frac{1}{n^2}
\]
We use the fact that $|\sin \left( \frac{x}{2} \right) | \geq \frac{|x|}{2\pi}$.

\[
\sum_{m: \left| m - \frac{2^r \tau \lambda_k}{2\pi} \right| \geq \ell} \frac{1}{4^r} \sin^2\left( \frac{2\pi m}{2^r} - \tau \lambda_k \right)/2\right)
\leq \sum_{m: \left| m - \frac{2^r \tau \lambda_k}{2\pi} \right| \geq \ell} \frac{1}{4|m - 2^{r-1} \tau \lambda_k / \pi|^2}
\leq 2 \times \frac{1}{4} \sum_{n=\ell}^{\infty} \frac{1}{n^2}
\leq \frac{1}{2(\ell - 1)} = O \left( \frac{1}{2^r \epsilon} \right).
\]
Summary: quantum phase estimation (textbook version)

A high level statement: QPE returns an energy estimate that is close to a random eigenvalue of $\tau_H$ with large probability.

A precise statement: QPE returns an energy estimate $\hat{\lambda}$ that is $\epsilon$-close to a random $\tau^k$ with probability at least $1 - O(\epsilon^{-1/2} - r)$, where $\hat{k} = k$ with probability $|c_k|^2$.

Runtime of the algorithm: apply control-e\(-i\tau_H^2r\) times.

$r = O(\epsilon^{-1})$ for constant success probability.

---

$^3$Nielsen and Chuang, *Quantum Computation and Quantum Information.*
A high level statement: QPE returns an energy estimate that is close to a random eigenvalue of $\tau H$ with large probability.
A high level statement: QPE returns an energy estimate that is close to a random eigenvalue of $\tau H$ with large probability.

A precise statement: QPE returns an energy estimate $\hat{\lambda}$ that is $\epsilon$-close to a random $\tau \lambda_{\hat{k}}$ with probability at least $1 - \mathcal{O}(\epsilon^{-1}2^{-r})$, where $\hat{k} = k$ with probability $|c_k|^2$.

---

$^3$Nielsen and Chuang, *Quantum Computation and Quantum Information.*
Summary: quantum phase estimation (textbook version)

- **A high level statement:** QPE returns an energy estimate that is close to a random eigenvalue of $\tau H$ with large probability.

- **A precise statement:** QPE returns an energy estimate $\hat{\lambda}$ that is $\epsilon$-close to a random $\tau \lambda_{\hat{k}}$ with probability at least $1 - O(\epsilon^{-1}2^{-r})$, where $\hat{k} = k$ with probability $|c_k|^2$.

- **Runtime of the algorithm:** apply control-$e^{-i\tau H}$ $2^r$ times. $2^r = O(\epsilon^{-1})$ for constant success probability.

---

3 Nielsen and Chuang, Quantum Computation and Quantum Information.
Use QPE for ground state energy

- Recall that the ground state energy is $\lambda_0$, the lowest eigenvalue of $H$. 
Use QPE for ground state energy

- Recall that the ground state energy is $\lambda_0$, the lowest eigenvalue of $H$.

- We need to ensure that the energy we get correspond to the ground state rather than excited states (ensure $\hat{k} = 0$): generate $O(1/|c_0|^2)$ samples for $\hat{\lambda}$ and take the minimum.
Use QPE for ground state energy

- Recall that the ground state energy is $\lambda_0$, the lowest eigenvalue of $H$.

- We need to ensure that the energy we get correspond to the ground state rather than excited states (ensure $\hat{k} = 0$): generate $O(1/|c_0|^2)$ samples for $\hat{\lambda}$ and take the minimum.

- Make sure in all of these samples the energy estimate is close to some eigenvalue. This happens with probability

$$\left(1 - O\left(\frac{1}{\epsilon 2^r}\right)\right)^{O(1/|c_0|^2)} = \Omega(1).$$
Use QPE for ground state energy

▶ Recall that the ground state energy is $\lambda_0$, the lowest eigenvalue of $H$.

▶ We need to ensure that the energy we get correspond to the ground state rather than excited states (ensure $\hat{k} = 0$): generate $\mathcal{O}(1/|c_0|^2)$ samples for $\hat{\lambda}$ and take the minimum.

▶ Make sure in all of these samples the energy estimate is close to some eigenvalue. This happens with probability

$$\left(1 - \mathcal{O}\left(\frac{1}{\epsilon 2^r}\right)\right)^{\mathcal{O}(1/|c_0|^2)} = \Omega(1).$$

▶ We therefore need $\epsilon 2^r = \Omega(1/|c_0|^2)$. 

Total cost (number of times to apply control-$e^{-i\tau H}$) for getting the ground state energy is

$$\frac{1}{|c_0|^2} \times \mathcal{O}\left(\frac{1}{\epsilon|c_0|^2}\right) = \mathcal{O}\left(\frac{1}{\epsilon|c_0|^4}\right).$$

$\# \text{ runs}$  $\mathcal{O}$  $\epsilon|c_0|^2$  $\epsilon|c_0|^4$.

We do not need to know $c_0$, but assume $|c_0|^2 \geq \eta$. Then the total cost is $\mathcal{O}(\epsilon^{-1} - \eta^{-2})$ (Heisenberg-limited scaling).

$r = \mathcal{O}(\log(\epsilon^{-1} - \eta^{-1}))$ is how deep the circuit is in each run. This is called the circuit depth. $r = \mathcal{O}(\log(\epsilon^{-1} - \eta^{-1}))$ is the number of ancilla qubits.
Total cost (number of times to apply control-$e^{-i\tau H}$) for getting the ground state energy is

$$\frac{1}{|c_0|^2} \times O\left(\frac{1}{\epsilon |c_0|^2}\right) = O\left(\frac{1}{\epsilon |c_0|^4}\right).$$

We do not need to know $c_0$, but assume $|c_0|^2 \geq \eta$. Then the total cost is $O(\epsilon^{-1}\eta^{-2})$ (Heisenberg-limited scaling).
Total cost (number of times to apply control-$e^{-i\tau H}$) for getting the ground state energy is
\[
\frac{1}{|c_0|^2} \times \mathcal{O}\left(\frac{1}{\epsilon|c_0|^2}\right) = \mathcal{O}\left(\frac{1}{\epsilon|c_0|^4}\right).
\]

We do not need to know $c_0$, but assume $|c_0|^2 \geq \eta$. Then the total cost is $\mathcal{O}(\epsilon^{-1}\eta^{-2})$ (Heisenberg-limited scaling).

$2^r = \mathcal{O}(\epsilon^{-1}\eta^{-1})$ is how deep the circuit is in each run. This is called the circuit depth. $r = \mathcal{O}(\log(\epsilon^{-1}\eta^{-1}))$ is the number of ancilla qubits.
We have looked at two algorithms for finding the ground state energy.
We have looked at two algorithms for finding the ground state energy.

- Textbook version QPE can estimate the ground state energy with an imperfect initial guess but requires many ancilla qubits.
We have looked at two algorithms for finding the ground state energy.

- Textbook version QPE can estimate the ground state energy with an imperfect initial guess but requires many ancilla qubits.
- Hadamard test can estimate the ground state energy using a single ancilla qubit, but requires a perfect initial guess.
We have looked at two algorithms for finding the ground state energy.

- Textbook version QPE can estimate the ground state energy with an imperfect initial guess but requires many ancilla qubits.
- Hadamard test can estimate the ground state energy using a single ancilla qubit, but requires a perfect initial guess.
- Both achieve the Heisenberg-limited scaling.

Coming up next: can we have all the good features?
The Hamiltonian $H$ has eigenvalues $\lambda_0, \lambda_1, \ldots$. For simplicity we assume $\text{spec}(H) \subset (-\pi/4, \pi/4)$.
The spectral density

- The Hamiltonian $H$ has eigenvalues $\lambda_0, \lambda_1, \cdots$. For simplicity we assume $\text{spec}(H) \subset (-\pi/4, \pi/4)$.

- An initial guess $|\Phi\rangle$ induces a probability distribution which we call the spectral density:

$$
\mu(x) = \sum_k p_k \delta(x - \lambda_k),
$$

where $p_k = |\langle \Psi_k | \Phi \rangle|^2$ and $\delta(\cdot)$ is the Dirac delta function.
The spectral density

- The Hamiltonian $H$ has eigenvalues $\lambda_0, \lambda_1, \cdots$. For simplicity we assume $\text{spec}(H) \subset (-\pi/4, \pi/4)$.

- An initial guess $|\Phi\rangle$ induces a probability distribution which we call the spectral density:

  \[ \mu(x) = \sum_k p_k \delta(x - \lambda_k), \]

  where $p_k = |\langle \Psi_k | \Phi \rangle|^2$ and $\delta(\cdot)$ is the Dirac delta function.

- If $X \sim \mu(x)$ then $X = \lambda_k$ w.p. $p_k$. Note that $\sum_k p_k = 1$. 

The spectral density

- The Hamiltonian $H$ has eigenvalues $\lambda_0, \lambda_1, \cdots$. For simplicity we assume $\text{spec}(H) \subset (-\pi/4, \pi/4)$.

- An initial guess $|\Phi\rangle$ induces a probability distribution which we call the spectral density:

  $$\mu(x) = \sum_k p_k \delta(x - \lambda_k),$$

  where $p_k = |\langle \Psi_k | \Phi \rangle|^2$ and $\delta(\cdot)$ is the Dirac delta function.

- If $X \sim \mu(x)$ then $X = \lambda_k$ w.p. $p_k$. Note that $\sum_k p_k = 1$.

- This distribution contains all the information about the spectrum.
\[ \langle \Phi | e^{-itH} | \Phi \rangle = \sum_k p_k e^{-it\lambda_k} = \int_{\mathbb{R}} \int_{\mathbb{T}} e^{-it\pi} \mu(x) dx. \]

This is the Fourier transform (either on \( \mathbb{R} \) or on the torus) of the distribution \( \mu(x) \).

\textbf{Figure:} The Hadamard test circuit.
The Hadamard test circuit outputs the expectation value

\[ \langle \Phi | e^{-itH} | \Phi \rangle = \sum_k p_k e^{-it\lambda_k} \]

\[ = \int_{-\pi}^{\pi} e^{-itx} \mu(x) dx. \]
The Hadamard test circuit outputs the expectation value

$$\langle \Phi | e^{-itH} | \Phi \rangle = \sum_k p_k e^{-it\lambda_k}$$

$$= \int_{-\pi}^{\pi} e^{-itx} \mu(x) dx.$$

This is the Fourier transform (either on $\mathbb{R}$ or on the torus) of the distribution $\mu(x)$. 

Figure: The Hadamard test circuit.
We can try recovering the spectrum density $\mu(x)$ through inverse Fourier transform. If we can then we will have the ground state energy.

Problem: we can only run the Hadamard test for a finite number of times, and the cost grows with $t$. $\mu(x)$ is a linear combination of Dirac delta functions. The Fourier coefficients do not decay.
We can try recovering the spectrum density $\mu(x)$ through inverse Fourier transform. If we can then we will have the ground state energy.

Problem: we can only run the Hadamard test for a finite number of times, and the cost grow with $t$. 
We can try recovering the spectrum density $\mu(x)$ through inverse Fourier transform. If we can then we will have the ground state energy.

**Problem:** we can only run the Hadamard test for a finite number of times, and the cost grow with $t$.

$\mu(x)$ is a linear combination of Dirac delta functions. The Fourier coefficients do not decay.
▶ We can instead compute a **smeared spectral density**:

\[ \tilde{\mu}(x) = \mu \ast M(x) = \int \mu(y)M(x - y)dy, \]

where \( M(\cdot) \) is a kernel (mollifier).
We can instead compute a smeared spectral density:

\[
\tilde{\mu}(x) = \mu \ast M(x) = \int \mu(y)M(x-y)\,dy,
\]

where \( M(\cdot) \) is a kernel (mollifier).

The Fourier coefficients are

\[
\hat{\mu}_k = \hat{\mu}_k \hat{M}_k,
\]

where \( \hat{\mu}_k = \langle \Phi | e^{-i k H} | \Phi \rangle \) can be computed from the Hadamard test, and \( \hat{M}_k \) is computed from \( M(x) \).
We can instead compute a smeared spectral density:

\[
\tilde{\mu}(x) = \mu \ast M(x) = \int \mu(y)M(x - y)dy,
\]

where \(M(\cdot)\) is a kernel (mollifier).

The Fourier coefficients are

\[
\hat{\mu}_k = \hat{\mu}_k \hat{M}_k,
\]

where \(\hat{\mu}_k = \langle \Phi | e^{-ikH} | \Phi \rangle\) can be computed from the Hadamard test, and \(\hat{M}_k\) is computed from \(M(x)\).

We want the kernel \(M(x)\) to be \(2\pi\)-periodic finite-degree trigonometric polynomial (the cost depend on the degree), and integrates to 1.
An example of such a kernel\textsuperscript{a}

\[ M(x) = \frac{1}{\mathcal{N}_{d,\delta}} T_d \left( 1 + 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)} \right), \]

where \( T_d \) is the \( d \)th Chebyshev polynomial of the first kind, and

\[ \mathcal{N}_{d,\delta} = \int_{-\pi}^{\pi} T_d \left( 1 + 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)} \right) dx. \]

\textsuperscript{a}Lin, Tong, 2020, *Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers.*
An example of such a kernel\textsuperscript{a}

\[
M(x) = \frac{1}{N_{d,\delta}} T_d \left(1 + 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)}\right),
\]

where \(T_d\) is the \(d\)th Chebyshev polynomial of the first kind, and

\[
N_{d,\delta} = \int_{-\pi}^{\pi} T_d \left(1 + 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)}\right) dx.
\]

\textsuperscript{a}Lin, Tong, 2020, Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers.

\textbf{Figure:} The kernel function \(M(x)\). The degree is \(O(\delta^{-1} \log(\epsilon^{-1}))\).
The quantum eigenvalue estimation problem (QEEP)\textsuperscript{a}

With $|\langle \Phi | \Psi_0 \rangle|^2 \geq \eta$ and to estimate ground state energy to precision $\epsilon$: total runtime $O(\epsilon^{-4}\eta^{-2})$, circuit depth $O(\epsilon^{-1})$, single ancilla, imperfect initial guess.

\textsuperscript{a}Somma, 2019, Quantum eigenvalue estimation via time series analysis.

Other relevant works:

- O'Brien et al., 2020, Error mitigation via verified phase estimation.
- Lu et al., 2020, Algorithms for Quantum Simulation at Finite Energies.
- Russo et al., 2020, Evaluating energy differences on a quantum computer with robust phase estimation.
The quantum eigenvalue estimation problem (QEEP)\(^a\)

With \( |\langle \Phi | \Psi_0 \rangle|^2 \geq \eta\) and to estimate ground state energy to precision \(\epsilon\): total runtime \(\mathcal{O}(\epsilon^{-4}\eta^{-2})\), circuit depth \(\mathcal{O}(\epsilon^{-1})\), single ancilla, imperfect initial guess.

\(\epsilon^{-4}\) comes from

- Number of terms in the Fourier expansion \(\mathcal{O}(\epsilon^{-1})\).
- To evaluate their sum to precision \(\mathcal{O}(\eta)\), we need \(\mathcal{O}(\epsilon^{-2}\eta^{-2})\) samples for each term.
- Average cost for each term is \(\mathcal{O}(\epsilon^{-1})\) (time needed for system to evolve).

\(^a\)Somma, 2019, *Quantum eigenvalue estimation via time series analysis.*
The quantum eigenvalue estimation problem (QEEP)\(^a\)

With \(|\langle \Phi | \Psi_0 \rangle|^2 \geq \eta\) and to estimate ground state energy to precision \(\epsilon\): total runtime \(\mathcal{O}(\epsilon^{-4} \eta^{-2})\), circuit depth \(\mathcal{O}(\epsilon^{-1})\), single ancilla, imperfect initial guess.

\(\epsilon^{-4}\) comes from

- Number of terms in the Fourier expansion \(\mathcal{O}(\epsilon^{-1})\).
- To evaluate their sum to precision \(\mathcal{O}(\eta)\), we need \(\mathcal{O}(\epsilon^{-2} \eta^{-2})\) samples for each term.
- Average cost for each term is \(\mathcal{O}(\epsilon^{-1})\) (time needed for system to evolve).

\(^a\)Somma, 2019, *Quantum eigenvalue estimation via time series analysis.*
The quantum eigenvalue estimation problem (QEEP)\(^a\)

With \(|\langle \Phi |\Psi_0 \rangle|^2 \geq \eta\) and to estimate ground state energy to precision \(\epsilon\): total runtime \(\mathcal{O}(\epsilon^{-4}\eta^{-2})\), circuit depth \(\mathcal{O}(\epsilon^{-1})\), single ancilla, imperfect initial guess.

\(\epsilon^{-4}\) comes from

- Number of terms in the Fourier expansion \(\mathcal{O}(\epsilon^{-1})\).
- To evaluate their sum to precision \(\mathcal{O}(\eta)\), we need \(\mathcal{O}(\epsilon^{-2}\eta^{-2})\) samples for each term.
- Average cost for each term is \(\mathcal{O}(\epsilon^{-1})\) (time needed for system to evolve).

\[^a\text{Somma, 2019, } Quantum eigenvalue estimation via time series analysis.}\]

Other relevant works:

- O’Brien et al., 2020, Error mitigation via verified phase estimation.
- Lu et al., 2020, Algorithms for Quantum Simulation at Finite Energies.
- Russo et al., 2020, Evaluating energy differences on a quantum computer with robust phase estimation.
Early fault-tolerant quantum algorithms

- We call such an algorithm an early fault-tolerant quantum algorithm
Early fault-tolerant quantum algorithms

- We call such an algorithm an early fault-tolerant quantum algorithm.
- It has the following nice features:
We call such an algorithm an early fault-tolerant quantum algorithm.

It has the following nice features:
- It uses only one ancilla qubit.
Early fault-tolerant quantum algorithms

- We call such an algorithm an **early fault-tolerant quantum algorithm**

- It has the following nice features:
  - It uses only one ancilla qubit.
  - The circuit depth is $O(\epsilon^{-1})$ rather than $O(\epsilon^{-1}\eta^{-1})$ (as in textbook QPE).
Early fault-tolerant quantum algorithms

- We call such an algorithm an **early fault-tolerant quantum algorithm**

- It has the following nice features:
  - It uses only one ancilla qubit.
  - The circuit depth is $\mathcal{O}(\epsilon^{-1})$ rather than $\mathcal{O}(\epsilon^{-1} \eta^{-1})$ (as in textbook QPE).
  - Some degree of noise-robustness.
We call such an algorithm an **early fault-tolerant quantum algorithm**

- It has the following nice features:
  - It uses only one ancilla qubit.
  - The circuit depth is $O(\epsilon^{-1})$ rather than $O(\epsilon^{-1} \eta^{-1})$ (as in textbook QPE).
  - Some degree of noise-robustness.

- Early fault-tolerant quantum algorithms are the ones that take the above factors into account.
Can we do better?

It may not be the best strategy to compute a smeared spectral density: each peak corresponds to an eigenvalue, but we do not know if there are eigenvalues to the left.

We should not evaluate each term and sum them up: $O(\epsilon - 1)$ terms and $O(\epsilon - 1)$ average time makes the cost at least $O(\epsilon - 2)$.

Not all terms are equally important.
Can we do better?

- It may not be the best strategy to compute a smeared spectral density: each peak corresponds to an eigenvalue, but we do not know if there are eigenvalues to the left.

- We should not evaluate each term and sum them up: $O(\epsilon^{-1})$ terms and $O(\epsilon^{-1})$ average time makes the cost at least $O(\epsilon^{-2})$.

- Not all terms are equally important.
Can we do better?

- It may not be the best strategy to compute a smeared spectral density: each peak corresponds to an eigenvalue, but we do not know if there are eigenvalues to the left.

- We should not evaluate each term and sum them up: $O(\epsilon^{-1})$ terms and $O(\epsilon^{-1})$ average time makes the cost at least $O(\epsilon^{-2})$. 
Can we do better?

- It may not be the best strategy to compute a smeared spectral density: each peak corresponds to an eigenvalue, but we do not know if there are eigenvalues to the left.

- We should not evaluate each term and sum them up: $O(\epsilon^{-1})$ terms and $O(\epsilon^{-1})$ average time makes the cost at least $O(\epsilon^{-2})$.

- Not all terms are equally important.
The cumulative distribution function

Define the cumulative distribution function (CDF):

\[ C(x) = \int_{-\pi}^{x} \mu(y) dy = \sum_{k} p_k H(x - \lambda_k) = (H \ast \mu)(x), \]

where \( H(\cdot) \) is the Heaviside function.
Define the cumulative distribution function (CDF):

\[ C(x) = \int_{-\pi}^{x} \mu(y) \, dy = \sum_{k} p_{k} H(x - \lambda_{k}) = (H * \mu)(x), \]

where \( H(\cdot) \) is the Heaviside function.

Every jump of this piecewise constant function correspond to an eigenvalue of \( H \). In particular, the first jump correspond to the ground-state energy.
Define the **cumulative distribution function (CDF):**

\[ C(x) = \int_{-\pi}^{x} \mu(y) dy = \sum_{k} p_k H(x - \lambda_k) = (H * \mu)(x), \]

where \( H(\cdot) \) is the Heaviside function.

Every jump of this piecewise constant function correspond to an eigenvalue of \( H \). In particular, the **first jump** correspond to the ground-state energy.

\( C(x) \) is a monotonously increasing function. We can therefore find the first jump by **binary search.**
The cumulative distribution function

- Define the **cumulative distribution function (CDF)**:

$$C(x) = \int_{-\pi}^{x} \mu(y) \, dy = \sum_{k} p_k H(x - \lambda_k) = (H \ast \mu)(x),$$

where $H(\cdot)$ is the Heaviside function.

- Every jump of this piecewise constant function correspond to an eigenvalue of $H$. In particular, the **first jump** correspond to the ground-state energy.

- $C(x)$ is a monotonously increasing function. We can therefore find the first jump by **binary search**.

- $C(x)$ can be approximated using Hadamard-test data.
Figure: Evaluating the CDF by sampling from the quantum circuit. Note that we do not need to re-sample for each point.
Figure: Zoom-in around the ground state energy
We can approximate $C(x)$ with a smooth function

$$\tilde{C}(x) = (M * C')(x) = (M * H * \mu)(x).$$
The approximate CDF

- We can approximate $C(x)$ with a smooth function

$$\tilde{C}(x) = (M \ast C')(x) = (M \ast H \ast \mu)(x).$$

- The approximation is in the sense

$$C(x - \delta) - \epsilon' \leq \tilde{C}(x) \leq C(x + \delta) + \epsilon',$$

which guarantees that we can find the first jump within error $\delta$. 


The approximate CDF

We can approximate $C(x)$ with a smooth function

$$\tilde{C}(x) = (M * C)(x) = (M * H * \mu)(x).$$

The approximation is in the sense

$$C(x - \delta) - \epsilon' \leq \tilde{C}(x) \leq C(x + \delta) + \epsilon',$$

which guarantees that we can find the first jump within error $\delta$.

We can evaluate $\tilde{C}(x)$ through its Fourier transform

$$\hat{\tilde{C}}_k = \hat{M}_k \hat{H}_k \hat{\mu}_k,$$

$k = -d, -d + 1, \cdots, d$, where $\hat{M}_k$ and $\hat{H}_k$ are known and $\hat{\mu}_k$ is computed from the Hadamard test.
 Importance sampling

▶ If we want to evaluate $\tilde{C}(x)$, then we need

$$\tilde{C}(x) = \frac{1}{2\pi} \sum_{k=-d}^{d} \tilde{C}_k e^{ikx} = \frac{1}{2\pi} \sum_{k=-d}^{d} \hat{M}_k \hat{H}_k \hat{\mu}_k e^{ikx}.$$
Importance sampling

If we want to evaluate $\tilde{C}(x)$, then we need

$$\tilde{C}(x) = \frac{1}{2\pi} \sum_{k=-d}^{d} \tilde{C}_k e^{ikx} = \frac{1}{2\pi} \sum_{k=-d}^{d} \hat{M}_k \hat{H}_k \hat{\mu}_k e^{ikx}.$$

If we compute each $\hat{\mu}_k$ ($k = -d, -d+1, \cdots, d$) individually, then each will require evolution time $O(|k|)$, which means the total evolution time is at least $O(d^2) = O(\epsilon_{gs}^{-2})$. Impossible to achieve the Heisenberg limit.
Importance sampling

- If we want to evaluate $\tilde{C}(x)$, then we need

$$
\tilde{C}(x) = \frac{1}{2\pi} \sum_{k=-d}^{d} \tilde{C}_k e^{ikx} = \frac{1}{2\pi} \sum_{k=-d}^{d} \hat{M}_k \hat{H}_k \hat{\mu}_k e^{ikx}.
$$

- If we compute each $\hat{\mu}_k$ ($k = -d, -d+1, \cdots, d$) individually, then each will require evolution time $O(|k|)$, which means the total evolution time is at least $O(d^2) = O(\epsilon_{gs}^{-2})$. Impossible to achieve the Heisenberg limit.

- Notice that

$$
\sum_{k=-d}^{d} |\hat{M}_k \hat{H}_k| \leq 2\pi \sum_{k=-d}^{d} |\hat{H}_k| \leq \sum_{k=-d}^{d} \frac{C}{1 + |k|} = O(\log(d)).
$$

Many terms are unimportant. We can use importance sampling.
Sample random variable $K$ with distribution

$$K = k \text{ w.p. } \frac{|\hat{M}_k \hat{H}_k|}{\sum_{k' = -d}^{d} |\hat{M}_{k'} \hat{H}_{k'}|}.$$ 

Then

$$\tilde{C}(x) = \left( \sum_{k' = -d}^{d} |\hat{M}_{k'} \hat{H}_{k'}| \right) \mathbb{E}[e^{i\theta K} \hat{\mu}_K].$$
Sample random variable $K$ with distribution

$$K = k \text{ w.p. } \frac{|\hat{M}_k \hat{H}_k|}{\sum_{k' = -d}^d |\hat{M}_{k'} \hat{H}_{k'}|}.$$ 

Then

$$\tilde{C}(x) = \left( \sum_{k' = -d}^d |\hat{M}_{k'} \hat{H}_{k'}| \right) \mathbb{E}[e^{i\theta K} \mu_K].$$

We can therefore randomly sample $K = k_1, k_2, \ldots, k_{N_s}$, compute $e^{i\theta k_l} \mu_{k_l}$, take average, and then multiply by $\sum_{k' = -d}^d |\hat{M}_{k'} \hat{H}_{k'}|$. 
Sample random variable $K$ with distribution

$$K = k \text{ w.p. } \frac{|\hat{M}_k \hat{H}_k|}{\sum_{k'=-d}^{d} |\hat{M}_{k'} \hat{H}_{k'}|}.$$ 

Then

$$\tilde{C}(x) = \left( \sum_{k'=-d}^{d} |\hat{M}_{k'} \hat{H}_{k'}| \right) \mathbb{E}[e^{i\theta_K \hat{\mu}_K}].$$

We can therefore randomly sample $K = k_1, k_2, \ldots, k_{N_s}$, compute $e^{i\theta_{Kl} \hat{\mu}_{kl}}$, take average, and then multiply by $\sum_{k'=-d}^{d} |\hat{M}_{k'} \hat{H}_{k'}|$. 

The variance is

$$\mathcal{O}\left(\log(d)^2\right) \frac{1}{N_s}.$$
Locating the first jump only requires precision $\eta$ (lower bound of the initial overlap $|\langle \Psi_0 | \Phi \rangle|^2$). Therefore we need

$$\frac{O(\log(d)^2)}{N_s} \leq \eta^2 \implies N_s = O(\log(d)^2 \eta^{-2}).$$
Locating the first jump only requires precision $\eta$ (lower bound of the initial overlap $|\langle \Psi_0 | \Phi \rangle|^2$). Therefore we need

$$\frac{\mathcal{O}(\log(d)^2)}{N_s} \leq \eta^2 \implies N_s = \mathcal{O}(\log(d)^2 \eta^{-2}).$$

If we want to compute the ground-state energy to precision $\epsilon_{gs}$, we need $d \sim \epsilon_{gs}^{-1}$. Then the total evolution time is

$$\mathcal{O}(d) \times N_s = \mathcal{O}(\epsilon_{gs}^{-1} \eta^{-2}).$$
Comparison with other ground-state energy algorithms:

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Runtime</th>
<th>Circ. depth</th>
<th>Ancilla qubits</th>
</tr>
</thead>
<tbody>
<tr>
<td>QPE</td>
<td>$\epsilon^{-1} \eta^{-2}$</td>
<td>$\epsilon^{-1} \eta^{-1}$</td>
<td>$\log(\epsilon^{-1} \eta^{-1})$</td>
</tr>
<tr>
<td>Modified QPE$^{4,5}$</td>
<td>$\epsilon^{-1} \eta^{-1}$</td>
<td>$\epsilon^{-1}$</td>
<td>$\log(\epsilon^{-1})$</td>
</tr>
<tr>
<td>Binary search$^6$</td>
<td>$\epsilon^{-1} \eta^{-1/2}$</td>
<td>$\epsilon^{-1}$</td>
<td>$\log(\eta^{-1})$</td>
</tr>
<tr>
<td>QEEP</td>
<td>$\epsilon^{-4} \eta^{-2}$</td>
<td>$\epsilon^{-1}$</td>
<td>1</td>
</tr>
<tr>
<td>CDF-based$^7$</td>
<td>$\epsilon^{-1} \eta^{-2}$</td>
<td>$\epsilon^{-1}$</td>
<td>1</td>
</tr>
</tbody>
</table>

---

$^4$Knill, Ortiz, Somma, 2006, *Optimal quantum measurements of expectation values of observables.*  
$^5$Berry, Su, Gyurik, et al., 2022, *Quantifying quantum advantage in topological data analysis.*  
Comparison with other ground-state energy algorithms:

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Runtime</th>
<th>Circ. depth</th>
<th>Ancilla qubits</th>
</tr>
</thead>
<tbody>
<tr>
<td>QPE</td>
<td>$\epsilon^{-1}\eta^{-2}$</td>
<td>$\epsilon^{-1}\eta^{-1}$</td>
<td>$\log(\epsilon^{-1}\eta^{-1})$</td>
</tr>
<tr>
<td>Modified QPE$^{4,5}$</td>
<td>$\epsilon^{-1}\eta^{-1}$</td>
<td>$\epsilon^{-1}$</td>
<td>$\log(\epsilon^{-1})$</td>
</tr>
<tr>
<td>Binary search$^6$</td>
<td>$\epsilon^{-1}\eta^{-1/2}$</td>
<td>$\epsilon^{-1}$</td>
<td>$\log(\eta^{-1})$</td>
</tr>
<tr>
<td>QEEP</td>
<td>$\epsilon^{-4}\eta^{-2}$</td>
<td>$\epsilon^{-1}$</td>
<td>1</td>
</tr>
<tr>
<td>CDF-based$^7$</td>
<td>$\epsilon^{-4}\eta^{-2}$</td>
<td>$\epsilon^{-1}$</td>
<td>1</td>
</tr>
</tbody>
</table>

This algorithm achieves **Heisenberg-limit scaling**, uses **single ancilla qubit**, and can use **imperfect initial guess**.

---

$^4$Knill, Ortiz, Somma, 2006, *Optimal quantum measurements of expectation values of observables.*

$^5$Berry, Su, Gyurik, et al., 2022, *Quantifying quantum advantage in topological data analysis.*


Extensions

- Computing ground state observable expectation values using a modified circuit.\(^8\)

---

\(^8\) Zhang, Wang, Johnson, 2021, *Computing ground state properties with early fault-tolerant quantum computers.*


\(^10\) Wang, Stilck-França, Zhang, Zhu, and Johnson, 2022, *Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision.*

\(^11\) Dong, Lin, **Tong,** 2022, *Ground state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices.*
Extensions

- Computing ground state observable expectation values using a modified circuit.\(^8\)

- Input model: \(e^{-i\tau H}\) using Trotter \(\rightarrow\) unbiased implementation through sampling.\(^9\)

---

\(^8\) Zhang, Wang, Johnson, 2021, *Computing ground state properties with early fault-tolerant quantum computers*.


\(^10\) Wang, Stilck-França, Zhang, Zhu, and Johnson, 2022, *Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision*.

Extensions

- Computing ground state observable expectation values using a modified circuit.\(^8\)

- Input model: \( e^{-i\tau H} \) using Trotter \(\rightarrow\) unbiased implementation through sampling.\(^9\)

- Lower circuit depth with large gap: \( \epsilon^{-1} \rightarrow \Delta^{-1} \).\(^{10}\)

---

\(^8\) Zhang, Wang, Johnson, 2021, *Computing ground state properties with early fault-tolerant quantum computers.*


\(^{10}\) Wang, Stilck-França, Zhang, Zhu, and Johnson, 2022, *Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision.*

Extensions

- Computing ground state observable expectation values using a modified circuit.\(^8\)

- Input model: \(e^{-i\tau H}\) using Trotter \(\rightarrow\) unbiased implementation through sampling.\(^9\)

- Lower circuit depth with large gap: \(\epsilon^{-1} \rightarrow \Delta^{-1}\).\(^{10}\)

- Adding \(Z\)-rotations (QSP-like) to get better runtime \(\tilde{O}(\epsilon^{-1}\eta^{-1})\).\(^{11}\)

---

\(^8\) Zhang, Wang, Johnson, 2021, *Computing ground state properties with early fault-tolerant quantum computers.*


\(^{10}\) Wang, Stilck-França, Zhang, Zhu, and Johnson, 2022, *Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision.*

\(^{11}\) Dong, Lin, Tong, 2022, *Ground state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices.*
Using least-squares fitting to get the dominant frequency. Better circuit depth with large overlap.\textsuperscript{12}

\begin{itemize}
  \item Ding, Lin, 2022, \textit{Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation}.
  \item Kshirsagar, Katabarwa, Johnson, 2022, \textit{On proving the robustness of algorithms for early fault-tolerant quantum computers}.
  \item Ding, Dong, \textbf{Tong}, Lin, 2023, \textit{Robust ground-state energy estimation under depolarizing noise}.
\end{itemize}
Using least-squares fitting to get the dominant frequency. Better circuit depth with large overlap.\textsuperscript{12}

Robustness under simple noise models.\textsuperscript{13,14}

\textsuperscript{12}Ding, Lin, 2022, Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation.

\textsuperscript{13}Kshirsagar, Katabarwa, Johnson, 2022, On proving the robustness of algorithms for early fault-tolerant quantum computers.

\textsuperscript{14}Ding, Dong, \textbf{Tong}, Lin, 2023, Robust ground-state energy estimation under depolarizing noise.
In the Hadamard test circuit, consider the following noise model: each time we apply control-$e^{-iH\tau}$, with probability $1 - e^{-\alpha\tau}$ the system decoheres into the maximally mixed state (loses all information).

---

Robustness under simple noise

- In the Hadamard test circuit, consider the following noise model: each time we apply control-$e^{-iH\tau}$, with probability $1 - e^{-\alpha\tau}$ the system decoheres into the maximally mixed state (loses all information).

- This is a reasonable model in random circuit.\(^{15}\)

\(^{15}\) Dalzell, Hunter-Jones, Brandao, 2021, *Random quantum circuits transform local noise into global white noise.*
Robustness under simple noise

- In the Hadamard test circuit, consider the following noise model: each time we apply control-$e^{-iH\tau}$, with probability $1 - e^{-\alpha\tau}$ the system decoheres into the maximally mixed state (loses all information).

- This is a reasonable model in random circuit.\(^{15}\)

- Then the output $\hat{m}$ (when measuring the ancilla qubit) will satisfy

$$\mathbb{E}[-1^{\hat{m}}] = e^{-\alpha t} \langle \Phi | e^{-iHt} | \Phi \rangle.$$ 

\(^{15}\)Dalzell, Hunter-Jones, Brandao, 2021, *Random quantum circuits transform local noise into global white noise.*
Robustness under simple noise

- In the Hadamard test circuit, consider the following noise model: each time we apply control-$e^{-iH\tau}$, with probability $1 - e^{-\alpha\tau}$ the system decoheres into the maximally mixed state (loses all information).

- This is a reasonable model in random circuit.\(^{15}\)

- Then the output $\hat{m}$ (when measuring the ancilla qubit) will satisfy

$$\mathbb{E}[-1^{\hat{m}}] = e^{-\alpha t} \langle \Phi | e^{-iHt} | \Phi \rangle .$$

- It is therefore possible to mitigate the error by multiplying $e^{\alpha t}$ (acceptable overhead if $\alpha$ not too large).

\(^{15}\)Dalzell, Hunter-Jones, Brandao, 2021, *Random quantum circuits transform local noise into global white noise.*
What happens to the textbook QPE under the same noise model?
What happens to the textbook QPE under the same noise model?

Maximally mixed state is invariant under any unitary operation. Therefore the ancilla register has $1 - e^{-\alpha^2 r \tau}$ probability of returning an uniformly random bit string.

\[
\Pr[E_{\text{meas, noisy}} = 2\pi m/2^r] = \Pr[E_{\text{meas}} = 2\pi m/2^r] e^{-\alpha^2 r \tau} + \frac{1 - e^{-\alpha^2 r \tau}}{2^r}.
\]
What happens to the textbook QPE under the same noise model?

Maximally mixed state is invariant under any unitary operation. Therefore the ancilla register has $1 - e^{-\alpha^2r\tau}$ probability of returning an uniformly random bit string.

$$
\Pr[E_{\text{meas, noisy}} = 2\pi m/2^r] = \Pr[E_{\text{meas}} = 2\pi m/2^r]e^{-\alpha^2r\tau} + \frac{1 - e^{-\alpha^2r\tau}}{2r}.
$$

Disastrous when taking the minimum (it is possible for any bit string to show up).
Experimental implementation

- A modified version of this algorithm is implemented using superconducting qubits.\textsuperscript{a}

- Active space with up to 4 spatial orbitals, 0.1 mHa

- Variational compilation, better Fourier coefficients,\textsuperscript{b} error mitigation.

\textsuperscript{a}Blunt, Caune, Izsák, Campbell, Holzmann, 2023, Statistical phase estimation and error mitigation on a superconducting quantum processor.

\textsuperscript{b}Wan, Berta, and Campbell, 2021, Randomized quantum algorithm for statistical phase estimation.
Compare with Kitaev’s iterative phase estimation algorithm.

Figure: O’Malley et al., 2016, *Scalable Quantum Simulation of Molecular Energies.*
For early fault-tolerant quantum computers, we may need to optimize for a number of metrics (circuit depth, number of qubits, error robustness) rather than just the runtime.
Conclusions

For early fault-tolerant quantum computers, we may need to optimize for a number of metrics \((\text{circuit depth, number of qubits, error robustness})\) rather than just the runtime.

It is possible to optimize for these metrics while keeping the Heisenberg-limited precision scaling.
Conclusions

- For early fault-tolerant quantum computers, we may need to optimize for a number of metrics (circuit depth, number of qubits, error robustness) rather than just the runtime.

- It is possible to optimize for these metrics while keeping the Heisenberg-limited precision scaling.

- We do not know what the first fault-tolerant quantum computer will be like, but the framework of early FTQC algorithms may be flexible enough to be useful.