Universal quantum algorithmic cooling on a quantum computer

Pei Zeng,† Jinzhao Sun,‡ and Xiao Yuan

1Pritzker School of Molecular Engineering, The University of Chicago, Illinois 60637, USA
2Center on Frontiers of Computing Studies, Peking University, Beijing 100871, China
3Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, UK
4School of Computer Science, Peking University, Beijing 100871, China
(Dated: June 6, 2022)

Quantum cooling, a deterministic process that drives any state to the lowest eigenstate, has been widely used from studying ground state properties of chemistry and condensed matter quantum physics to general optimization problems. However, the cooling procedure is non-unitary, making its practical realization on a quantum computer a non-trivial task. Here, we propose universal quantum cooling algorithms with small resource costs. By utilizing a dual-phase representation of decaying functions, we show how to universally and deterministically realize a general cooling procedure with shallow quantum circuits. We demonstrate its applications in cooling an arbitrary input state with known ground state energy, corresponding to satisfactory, linear algebra tasks, and quantum state compiling tasks, and preparing eigenstates with unknown eigenenergies, corresponding to quantum many-body problems. Compared to quantum phase estimation, our method uses only one ancillary qubit, showing exponential improvement of the circuit complexity with respect to the final state infidelity while maintaining the Heisenberg limit of the eigenenergy estimation with proper cooling functions. We numerically benchmark the algorithms for the 8-qubit Heisenberg model and verify its feasibility for accurately finding eigenenergies and obtaining eigenstate measurements. Our work paves the way for efficient and universal quantum algorithmic cooling with near-term and universal fault-tolerant quantum devices.

Introduction.— Many tasks, whether finding ground states [1, 2], studying low temperature physics [3, 4], solving optimization problems [5, 6], or even simulating quantum dynamics [7], can be abstracted as a cooling problem. Among the various task-specific classical algorithms — such as perturbation theories [8], variational approaches [9, 10], self-consistent embedding methods [11, 12], machine learning [13, 14], etc — imaginary time evolution [15] defines a natural and universal cooling procedure. Consider a Hamiltonian \( H \), the imaginary time evolution \( e^{-\tau H} \) with a real-valued time \( \tau \) monotonically lower down the average energy and deterministically drives an arbitrary pure state to the lowest eigenstate that has a nonzero overlap. Despite being mathematically universal, realizing imaginary time evolution for an arbitrary Hamiltonian is by no means an easy problem. The notorious sign problem [16] in Monte Carlo implementation of imaginary time evolution has limited its usage for solving general strongly correlated problems [17, 18].

Can we realize a cooling procedure more efficiently on a quantum computer? The answer is yes! The most prominent algorithm is quantum phase estimation [19, 20], which efficiently projects an arbitrary input quantum state to an eigenstate of the Hamiltonian, given the input state has a nonvanishing overlap with the target eigenstate. However, quantum phase estimation generally requires a deep circuit that is only realizable with a universal quantum computer [21, 22]. With near-term noisy intermediate-scale quantum devices, variational approaches have been developed to emulate the imaginary time evolution with fixed or adaptive parametrized quantum circuits [23–27]. Nevertheless, whether such circuits exist or how to efficiently find them still remains unclear [28–30]. Several recent works have indicated the possibility of more efficient cooling algorithms [31–34], whereas they still assume the usage of many ancillary qubits or unconventional computational models.

Here, we propose a new paradigm for simulating general quantum cooling procedures on a quantum computer. For a general decaying function \( g(h) \), such as \( g = e^{-|h|} \) or \( g = e^{-h^2} \), we show how to realize a cooling procedure \( g(\tau H) \) for Hamiltonian \( H \) using solely real-time evolution and proper classical post-processing. The algorithm essentially implements general non-Hermitian dynamics, and it can efficiently find energy spectra and corresponding eigenstates, for any initial states that have nonvanishing overlaps with the target state. Compared to quantum phase estimation, our method only needs an exponentially shallower circuit with only one ancillary qubit and maintains the Heisenberg limit of the eigenenergy estimations when using proper cooling functions. Compared to variational approaches, our algorithm works deterministically without any assumption on parametrized circuits.

Quantum algorithmic cooling.— We start with a real-valued function \( g(h) \) satisfying strictly non-increasing absolute value, \( |g(h′)| < |g(h)| \), \( \forall h′ > h > 0 \) or \( h′ < h < 0 \) [35], and vanishing asymptotic value, \( \lim_{\tau \to \infty} |g(\tau h)|/g(\tau h)| = 0 \), \( \forall h′ > h > 0 \) or \( h′ < h < 0 \) or
Consider a positive Hamiltonian $H = \sum_i E_i |u_i\rangle \langle u_i|$ with eigenstates $\{ |u_i\rangle \}$ and eigenenergies $E_i \geq 0$ [36], it defines a cooling procedure $g(\tau H)$ with Hamiltonian $H$ and real-valued time $\tau$ for an arbitrary input state $|\psi_0\rangle = \sum_i c_i |u_i\rangle$ as

$$\langle \psi(\tau) \rangle = \frac{g(\tau H) |\psi_0\rangle}{\|g(\tau H) |\psi_0\rangle\|} = \sum_i p_i g(\tau E_i)^2 c_i |u_i\rangle,$$

where $p_i = |c_i|^2$. Since for larger eigenenergies the function $g(\tau E_i)$ decreases faster with $\tau$, the amplitudes of the normalized state $\langle \psi(\tau) \rangle$ concentrate to lower eigenstates, and the evolved state asymptotically approximates the lowest eigenstate $|u_0\rangle$ with nonzero $c_i$ for sufficiently large $\tau$.

To realize the cooling function $g(h)$ with valid Fourier transforms, we consider the dual form of $g(h)$ via its Fourier transform $g(h) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{ixh} dx$, with the inverse transform $f(x) = \int_{-\infty}^{\infty} g(h) e^{-ixh} dh$. Given the norm of $f(x)$ defined as $\|f\| = \int_{-\infty}^{\infty} |f(x)| dx$, we consider the normalized function $p(x) = |f(x)| / \|f\|$ and we have

$$g(h) = c \int_{-\infty}^{\infty} e^{ihx} p(x) e^{ixh} dx,$$

where $c = \|f\| / 2\pi$ and $e^{ihx} = f(x) / |f(x)|$. The function $g(h)$ is realizable if the following conditions hold.

$C1$: The normalization $\|f\|$ or $c$ is finite.

$C2$: $|1 - \int_{-L(c)}^{L(c)} p(x) dx| \leq \epsilon, \forall \epsilon \geq 0, \exists L(c) = O(\text{poly}(\frac{1}{\epsilon})).$

The first condition implies a finite norm of $\|f\|$ and the second condition implies that a finite frequency in $[-L(c), L(c)]$ is sufficient for approximating the normalized function $g(h)/c$, i.e., $|g(h)/c - \int_{-L(c)}^{L(c)} p(x) e^{ihx} dx| \leq \epsilon$. In Table I, we give several examples of $g(h)$. Since our examples have zero phase, for a simpler presentation, we set the phase terms to be 0 in the following discussion. Although, our following discussion naturally generalizes to the case with nonzero phases. The detailed analysis of the cooling functions is given in Supplementary Materials.

With the normalized dual phase representation $p(x)$ of $g(h)$, the cooling procedure of Eq. (1) becomes

$$\langle \psi(\tau) \rangle \propto \int_{-\infty}^{\infty} dp(x) |\phi(x\tau)\rangle,$$

which is now a superposition of real-time evolved states $|\phi(x\tau)\rangle = e^{ixH\tau} |\psi_0\rangle$ with amplitudes $dp(x) = p(x) dx$.

The linear-combination-of-unitary algorithm [33, 37, 38] could be exploited to create the superposition. However, creating such continuous-variable states is challenging with a digital quantum computer. Here, we propose an experimentally more feasible and efficient way to effectively realize the superposition.

| Type | $g(h)$ | $f(x)$ | $C1$ | $C2$
|------|--------|--------|------|------|
| Rectangular | $\eta(|h| \leq 1/2)$ | $\sin(x/2\pi)$ | $\infty$ | $\infty$
| Triangle | $(1 - |h|/|h| \leq 1)$ | $\sin^2(x/2\pi)$ | $2\pi$ | $6/\epsilon$
| Exponential | $e^{-|x|}$ | $\frac{2}{\pi e} + x$ | $1$ | $2/\epsilon$
| Gaussian | $e^{-h^2}$ | $\sqrt{\pi e}e^{-x^2/4}$ | $1$ | $2\sqrt{\ln(1/\epsilon)}$
| Hyperbolic | sech($h$) | $\pi \text{sech}(\pi x/2)$ | $\frac{2}{\pi \ln(1/\epsilon)}$

Instead of preparing the quantum state of Eq. (3), we focus on the goal to obtain observable arbitrary expectation values. Specifically, we aim to measure any observable $O$ of the evolved state $|\psi(\tau)\rangle$, i.e.,

$$\langle O \rangle_{\psi(\tau)} = \langle \psi(\tau) | O | \psi(\tau) \rangle = \frac{N_{\tau}(O)}{D_{\tau}},$$

where $N_{\tau}(O) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dp(x, x') \langle \phi(x\tau') | O | \phi(x\tau) \rangle$ and $D_{\tau} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dp(x, x') \langle \phi(x\tau') | \phi(x\tau) \rangle$ with $dp(x, x') = p(x) p(x') dx dx'$. Here $D_{\tau}$ can be simplified as $D_{\tau} = \int_{-\infty}^{\infty} dp(y) \langle \psi_0 | e^{i\tau H} | \psi_0 \rangle$ with $dp(y) = \tilde{p}(y) dy$. Since $N_{\tau}(O)$ is real, we can efficiently obtain it by sampling the distribution $dp(x, x')$ and then estimating the mean value $E_{x, x'} \langle \phi(x\tau') | O | \phi(x\tau) \rangle$, where each term is realizable on a quantum computer with the Hadamard test circuit [39].

Applications.—Here, we consider applications of quantum algorithmic cooling for quantum systems with known and unknown eigenenergies. The detailed algorithms can be found in Supplementary Materials.

First, we consider Hamiltonians $H$ with a known ground state energy $E_0 \geq 0$, corresponding to satisfactory problems [5, 6], linear algebra tasks [40–42], quantum state compiling [43–46], quantum error correction [47, 48], etc. For an initial state $|\psi_0\rangle = \sum_i c_i |u_i\rangle$ with a nonvanishing $p_0 = |\alpha_0|^2$, we can obtain the average of any observable of a well approximated the ground state $|\psi_0\rangle$ according to Eq. (4). However, when $E_0 > 0$, the denominator $D_{\tau} = \langle \psi_0 | g(\tau H) | \psi_0 \rangle / c^2$, may decrease exponentially with $\tau$, where $c = (2\pi/\|f\|)^{1/2}$ is
the exponential decay is to shift the Hamiltonian by $-E_0$, so that the shifted ground state energy is 0, and for large $\tau$, the denominator $D_\tau$ scales as $p_\tau$. Furthermore, it is not hard to see that for the valid cooling functions considered in Table I, the algorithm works for any eigenstate $|\epsilon_i\rangle$ with known eigenenergy $E_i$, by shifting the Hamiltonian by $-E_i$. The algorithm is efficient as long as the probability $p_\tau$ is not too small.

For the second task, we consider problems with unknown (inaccurate) eigenenergy, such as chemistry or condensed matter problems [50–53]. Starting with a known (inaccurate) eigenenergy, such as chemistry or condensed matter problems [50–53]. Starting with a guess $E$ of an eigenenergy (for instance, obtained from classical methods), we shift the Hamiltonian by $-E$ and estimate the denominator $D_\tau(\psi,\psi)$ as $D_\tau(\psi,\psi) = \langle \psi|g((H - E)\tau)^2|\psi\rangle/c^2$. Note that $D_\tau(\psi,\psi)$ is locally maximized to $p_\tau/c^2$ when $E = E_i$ is an eigenenergy of $H$, otherwise it exponentially decays with $\tau$. Then we can scan the denominator $D_\tau(\psi,\psi)$ with different trial energy $E$ and find the eigenenergies of $H$ via the peaks of $D_\tau(\psi,\psi)$. Furthermore, the denominator $D_\tau(\psi,\psi)$ is

$$D_\tau(\psi,\psi) = \int_{-\infty}^{\infty} d\tilde{p}(y)e^{-i\tau E}\langle \psi|e^{iy\tau H} |\psi\rangle,$$

where $\langle \psi|e^{iy\tau H} |\psi\rangle$ is obtained from quantum computers, and the information of $E$ and the distribution $d\tilde{p}(y)$ only appears as the coefficient. Therefore $D_\tau(\psi,\psi)$ with different $E$ could be obtained from different classical post-processing of the same measurement outcome $\langle \psi|e^{iy\tau H} |\psi\rangle$, making the sample-efficiency of the algorithm independent of the value and the number of trial energies $E$. Due to the local monotonic feature of $g$, there also exist efficient classical algorithms, such as binary search algorithms, to find the peaks.

In the above discussion, we have assumed a ‘good’ initial state that has a nonvanishing overlap with the target state, an assumption usually adopted in quantum phase estimation [54] or other non-variational ground-state preparation algorithms [31, 55, 56]. Although this assumption is hard to promise for arbitrary problems, it is reasonable in many practical scenarios. For example, we could just use a classical approximated solution [9, 10], such as the Hartree-Fock state or a tensor network solution, or we can apply adiabatic state preparation to enhance the state overlap [1, 2, 57].

Besides, one can use shallow parametrized quantum circuits and apply variational optimization to search for an approximated solution [58–60]. Integrating these approximation algorithms with our quantum algorithmic cooling, it provides an efficient way to study eigenstates and eigenenergies of many-body quantum systems.

**Error and resource estimation.**—When applying the cooling algorithm in practice, we have to take the error caused by finite circuit depth and finite sampling number into consideration. There are three dominant factors limiting the precision of our estimation — the finite imaginary time $\tau$, the cutoff $x_m$ of the integration, and the finite sampling number $N_M$ during the experiment. We summarize how those factors affect observable and eigenenergy estimation and refer to Supplementary Materials for details.

First, we consider the error of observable estimation with a known target eigenenergy. Denote the ideal (estimated) measurements with infinite (finite) $\tau$, $x_m$, and $N_M$ as $\langle O \rangle_i^{(\tau)}$ and $\langle O \rangle_i^{(\tau, x_m)}$, respectively. The effect of the three factors on observable estimation is given as follows.

**Theorem 1** (Accuracy of the observable estimation). Given constant $K > 0$, error $\varepsilon \in (0, 1)$, finite imaginary time $\tau \geq g^{-1}(\varepsilon p_j/6)/\Delta$, normalized cutoff time $x_m \geq \sqrt{2L}(\varepsilon p_j/12)$, and sample number $N_M \geq K/(\varepsilon p_j/6)^2$, the error between the expectation value estimation $\langle O \rangle_i^{(\tau, x_m)}$ and the ideal expectation value $\langle O \rangle_i^{(\tau)}$ is bounded by $|\langle O \rangle_i^{(\tau, x_m)} - \langle O \rangle_i^{(\tau)}| \leq \varepsilon|\langle O \rangle_i|$, with a failure probability of $\delta = 4\exp(-K/8)$. Here $|\langle O \rangle_i| = \sum_o |o|$, where $\{o\}$ are the Pauli coefficients $O = \sum_o o_1 P_1$, $\Delta = \min(|E_{j-1} - E_j|, |E_{j+1} - E_j|)$, and $p_j$ is the overlap with the target eigenstate.

**Theorem 4** states that the asymptotic time complexity of the cooling algorithm is determined by $g^{-1}(\varepsilon)L(\varepsilon)$ of the cooling function. The valid cooling functions in Table I all satisfy $g^{-1}(\varepsilon) = O(\log(1/\varepsilon))$. Therefore, we can choose $\tau$ to be $O(\log(1/(p_j\varepsilon)))$, which is logarithmic to the inverse state overlap $1/p_j$ and inverse error $1/\varepsilon$, and linear to the inverse gap $\Delta^{-1}$. For the cutoff $x_m$, we have shown that for the triangle and exponential cooling functions, we have $L(\varepsilon) = O(\text{poly}(1/\varepsilon))$, and for the Gaussian and secant hyperbolic cooling functions, we can achieve even better results $O(\sqrt{\log(1/\varepsilon)})$ or $L(\varepsilon) = O(\log(1/\varepsilon))$. Take the Gaussian function as an example, we have $x_m = O(\log(1/(\varepsilon p_j)))$. Note that the maximum real-time evolution is given by $t_m = \tau x_m$. As a result, we have the following result.

**Proposition 1.** The time (circuit) $t_m$ and sample $N_M$ complexity for the Gaussian cooling function is

$$t_m \sim O(\Delta^{-1} \log(1/(\varepsilon p_j))),$$

$$N_M \sim O(1/(\varepsilon p_j^2)).$$

The time or circuit complexity is logarithmic to $1/(p_j\varepsilon)$, which is exponentially better than that of quantum phase estimation or adiabatic state preparation, which is generally polynomial to $1/(p_j\varepsilon)$. The sample complexity is slightly worse than that of quantum phase estimation, which is $O(1/(p_j^2\varepsilon^2))$.

Next, we consider the error of eigenenergy estimation.

**Theorem 2** (Accuracy of the eigenenergy estimation). Given constant $K > 0$, error $\varepsilon \in (0, 1)$,
finite imaginary time \( \tau \geq 1/\kappa \), normalized cutoff time \( x_m \geq \sqrt{2L} ((1 - g(1)) p_j /4) \), and sample number \( N_M \geq 2 K p_j^{-1} (1 - g(1)) \), the error between the estimated eigenenergy \( \hat{E}_j \) and the ideal eigenenergy \( E_j \) is \( |\hat{E}_j - E_j| \leq \kappa \), with a failure probability of \( 2\delta = 4 \exp (-K/8) \).

Note that since only the imaginary time \( \tau \) depends on the inverse accuracy \( 1/\kappa \), the total cost \( \tau \cdot x_m \cdot N_M \) (circuit complexity \( \times \) sample complexity) also scales as \( 1/\kappa \), indicating Heisenberg’s limit for eigenenergy estimation, similar to a former result by Lin and Tong [61]. We refer to Table I and Table II in [39] for the complexity comparison in the eigenenergy estimation and observable estimation, respectively.

**Numerical test.**—Here we show numerical implementation of the algorithmic cooling method for the anisotropic Heisenberg Hamiltonian, 
\[
H = J \sum_i (\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + 2 \sigma_i^z \sigma_{i+1}^z) + h \sum_i \sigma_i^z,
\]
where \( J = 1 \) is the exchange coupling, \( \sigma_i^\alpha (\alpha = x, y, z) \) is the Pauli operator on the \( i \)th site, \( h = 1 \) is strength of a uniform magnetic field in the \( z \) direction, and we impose periodic boundary condition. We consider the initial state in the computation basis as \( |01010101\rangle \), which is close to the ground state and has nonzero overlap with a relatively small number of low-energy eigenstates. We first consider determining energy spectra by searching for the peaks of \( D_\tau(E) \) with the Gaussian cooling function. Using a finite imaginary time \( \tau = 1.7 \) and cutoff \( x_m = 4.4 \) according to \( \tau, x_m \propto \sqrt{2\log(1/\epsilon)} \) with error \( \epsilon \), and \( 10^5 \) number of samples for the integral, we show the energy spectra in Fig. 1. The maximum error introduced is below 0.01, which aligns with our error analysis. We further show the searching process with different time \( \tau \) in Ref. [39].

Next, we show the error dependence of the eigenstate preparation with a special focus on the time or circuit complexity. Suppose we aim to find the second eigenstate \( |u_2\rangle \), which has the largest overlap with the initial state. Given the associated eigenenergy \( E_2 \) found by the above method, we now analyze the error introduced from finite imaginary time and cutoff. Here, we focus on the state infidelity of the normalized state after cooling and the target ideal eigenstate, which can be expressed as \( \epsilon = 1 - \langle O \rangle_\tau \) with \( O = |u_2\rangle \langle u_2| \). We show the error dependence on the imaginary time \( \tau \) and the maximal total evolution time \( t_m = \tau x_m \) with Gaussian and exponential cooling functions in Fig. 2(a1-a2) and (b1-b2), respectively. We plot the theoretical upper bound of the infidelity in the dashed line in the figures (assuming finite \( \tau \), infinite \( x_m \), and infinite samples \( N_M \) and realistic infidelity with a constant resources using dots. In Fig 2(a1) and (b1), we consider finite \( \tau \) and \( x_m \) with \( 10^5 \) samples for calculating the integral, but ignore measurement errors. We see the exponential scaling with respect to the total evolution time \( t_m \) for the Gaussian cooling function, while it asymptotically becomes \( 1/\epsilon \) for the exponential cooling function. In Fig 2(a2) and (b2), we further consider single-shot measurement outcomes again with a total of \( 10^5 \) samples and show the error.

**FIG. 1.** Energy spectra search of the 8-site anisotropic Heisenberg model using the Gaussian cooling function. The solid line shows the denominator of \( D_\tau \) with \( \tau = 1.7 \) and cutoff \( x_m = 4.4 \). The dashed line shows the exact eigenenergy of the Heisenberg model. The figure inset shows the error of \( D_\tau \) with respect to that under the exact cooling. We set \( N_s = 10^5 \) in the Monte Carlo sampling of the integral, and we use the expectation value for each sample that ignores measurement shot noise.

**FIG. 2.** Error dependence of the total evolution time with (a1-a2) Gaussian and (b1-b2) exponential cooling functions. In (a1) and (b1), we adjust both the imaginary time \( \tau \) and the cutoff \( x_m \), and study the infidelity versus maximal total time \( t_m = \tau x_m \). We ignore the measurement shot noise in (a1, b1). In (a2) and (b2), we further consider measurement shot noise, which introduces more instability of the infidelity. Nevertheless, we still observe the exponential and linear inverse decay of the infidelity for the Gaussian and exponential functions, respectively. The errorbar is the standard deviation of \( D_\tau \) error over 10 independent repetitions of the entire setup. We use \( 10^5 \) samples in the Monte Carlo sampling of the integral. Dashed line: theoretical upper bound of the infidelity assuming infinite sample; Dots with error bar: realistic infidelity with a constant number of samples.
dependence accordingly.

**Conclusion & Discussion.**—We propose a universal and economic quantum algorithm to implement general cooling procedures. The algorithm is applicable to estimate eigenenergies and prepare eigenstates of problems with known and unknown eigenenergies when the overlap with the initial state is nonvanishing. In practice, we can use adiabatic state preparation or variational quantum eigensolver to prepare a good initial state. Our method is more competitive to quantum phase estimation or bare variational algorithms when the circuit depth and qubits are limited.

Another potential application is to study system properties under the finite-temperature condition. Given the ability of implementing general cooling function or preparing eigenstates, one may combine existing technique, such as Metropolis sampling [32], to effectively prepare thermal states and study finite-temperature physics. Besides, the algorithm may also help realize a general Markovian quantum process, which is described by a non-Hermitian Hamiltonian [62]. We can decompose the Hamiltonian to the real and imaginary part, and further expand the imaginary part using Fourier transform. In this way, we could simulate a general non-Hermitian dynamic process.

In this work, we mainly explore the power of continuous Fourier transform in Eq. (2) to construct the cooling function $g(h)$. We remark that the Fourier transform is a universal tool to construct any non-unitary process else than cooling. For example, one can create a “two-peak” cooling function to prepare a cat-like state, i.e., the superposition of two remote eigenstates. Meanwhile, we can explore different linear transforms of continuous or discrete functions, such as discrete Fourier transform and continuous wavelet transform, to study their possible usage in realizing even more general non-unitary processes. Several related recent ideas include Refs. [31–34, 61]. Comparison and combination of these ideas are also an interesting future work.

**ACKNOWLEDGMENTS**

This work is supported by the National Natural Science Foundation of China Grant No. 12175003. The numerics is supported by High-performance Computing Platform of Peking University.

*Note added.*—Recently, we became aware of a relevant work by Huo and Li [63], who considered a specific cooling function using the Fourier transform of the Lorentz-Gaussian function. They studied its application for realizing imaginary time evolution and discussed its combination with error mitigation for implementation with near-term quantum devices.

[1] D. S. Abrams and S. Lloyd, Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors, Phys. Rev. Lett. **83**, 5162 (1999).
[2] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Simulated quantum computation of molecular energies, Science **309**, 1704 (2005), https://www.science.org/doi/pdf/10.1126/science.1113479.
[3] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete, Quantum metrology sampling, Nature **471**, 87 (2011).
[4] M.-H. Yung and A. Aspuru-Guzik, A quantum–quantum metrology algorithm, Proceedings of the National Academy of Sciences **109**, 754 (2012), https://www.pnas.org/content/109/3/754.full.pdf.
[5] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm (2014), arXiv:1411.4028 [quant-ph].
[6] L Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices, Phys. Rev. X **10**, 021067 (2020).
[7] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, Polynomial-time quantum algorithm for the simulation of chemical dynamics, Proceedings of the National Academy of Sciences **105**, 18681 (2008), https://www.pnas.org/content/105/48/18681.full.pdf.
[8] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of quantum field theory in statistical physics* (Courier Corporation, 2012).
[9] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Courier Corporation, 2012).
[10] R. Orús, Tensor networks for complex quantum systems, Nature Reviews Physics **1**, 538 (2019).
[11] D. Sholl and J. A. Steekel, *Density Functional Theory: a Practical Introduction* (John Wiley & Sons, 2011).
[12] G. Knizia and G. K.-L. Chan, Density matrix embedding: A simple alternative to dynamical mean-field theory, Phys. Rev. Lett. **109**, 186404 (2012).
[13] J. Hermann, Z. Schätzle, and F. Noé, Deep-quantum network solution of the electronic schrödinger equation, Nature Chemistry **12**, 891 (2020).
[14] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, Ab initio solution of the many-electron schrödinger equation with deep neural networks, Phys. Rev. Research **2**, 033429 (2020).
[15] N. Metropolis and S. Ulam, The monte carlo method, Journal of the American Statistical Association **44**, 335 (1949), pMID: 18139350.
[16] M. Troyer and U.-J. Wiese, Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations, Phys. Rev. Lett. **94**, 170201 (2005).
[17] J. Kolorenc and L. Mitas, Applications of quantum monte carlo methods in condensed systems, Reports on Progress in Physics **74**, 026502 (2011).
[18] B. M. Austin, D. Y. Zubarev, and W. A. Lester, Quantum monte carlo and related approaches, Chemical Reviews **112**, 263 (2012).
[19] A. Y. Kitaev, Quantum measurements and the abelian stabilizer problem, arXiv preprint quant-ph/9511026 (1995).
[20] A. Y. Kitaev, A. Shen, M. N. Vyalyi, and M. N. Vyalyi, Classical and quantum computation, 47 (American Mathematical Soc., 2002).
[21] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven, Encoding electronic spectra in quantum circuits with linear t complexity, Phys. Rev. X 8, 041015 (2018).
[22] C. Gidney and M. Elkerä, How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits, Quantum 5, 433 (2021).
[23] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, Variational ansatz-based quantum simulation of imaginary time evolution, npj Quantum Information 5, 75 (2019).
[24] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, Theory of variational quantum simulation, Quantum 3, 191 (2019).
[25] M. Motta, C. Sun, A. T. K. Tan, M. J. O’Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nature Physics 16, 205 (2020).
[26] H. Nishi, T. Kosugi, and Y.-i. Matsushita, Implementation of quantum imaginary-time evolution method on nisq devices by introducing nonlocal approximation, npj Quantum Information 7, 85 (2021).
[27] D. Amaro, C. Modica, M. Rosenkranz, M. Fiorentini, M. Benedetti, and M. Lubasch, Filtering variational quantum algorithms for combinatorial optimization, Quantum Science and Technology 7, 015021 (2022).
[28] A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner, The power of quantum neural networks, Nature Computational Science 1, 403 (2021).
[29] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus (2021), arXiv:2101.02138 [quant-ph].
[30] Z. Holmes, A. Arrasmith, B. Yan, P. J. Coles, A. Albrecht, and A. T. Sornborger, Barren plateaus preclude learning scramblers, Phys. Rev. Lett. 126, 190501 (2021).
[31] Y. Ge, J. Tura, and J. I. Cirac, Faster ground state preparation and high-precision ground energy estimation with fewer qubits, Journal of Mathematical Physics 60, 022202 (2019), https://doi.org/10.1063/1.5027484.
[32] S. Lu, M. C. Bañuls, and J. I. Cirac, Algorithms for quantum simulation at finite energies, PRX Quantum 2, 020321 (2021).
[33] D.-B. Zhang, G.-Q. Zhang, Z.-Y. Xue, S.-L. Zhu, and Z. D. Wang, Continuous-variable assisted thermal quantum simulation, Phys. Rev. Lett. 127, 020502 (2021).
[34] K. Choi, D. Lee, J. Bonitati, Z. Qian, and J. Watkins, Rodeo algorithm for quantum computing, Phys. Rev. Lett. 127, 040505 (2021).
[35] The framework also works even if the function is not strictly non-increasing.
[36] For any finite dimensional Hamiltonian, we can always shift it by a constant value to make it positive.
[37] G. L. Long, Duality quantum computing and duality quantum information processing, International Journal of Theoretical Physics 50, 1305 (2011).
[38] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Simulating hamiltonian dynamics with a truncated taylor series, Phys. Rev. Lett. 114, 090502 (2015).
[39] See Supplementary Materials for detailed discussions on the cooling function and resource analysis.
[40] C. Bravo-Prieto, R. LaRose, M. Cerezo, Y. Subasi, L. Cincio, and P. J. Coles, Variational quantum linear solver: A hybrid algorithm for linear systems, arXiv preprint arXiv:1909.05820 (2019).
[41] X. Xu, J. Sun, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, Variational algorithms for linear algebra, Science Bulletin https://doi.org/10.1016/j.scib.2021.06.023 (2021).
[42] H.-Y. Huang, K. Bharti, and P. Rebentrost, Near-term quantum algorithms for linear systems of equations, arXiv preprint arXiv:1909.07344 (2019).
[43] S. Khatri, R. LaRose, A. Poremba, L. Cincio, A. T. Sornborger, and P. J. Coles, Quantum-assisted quantum compiling, Quantum 3, 140 (2019).
[44] K. Heya, Y. Suzuki, Y. Nakamura, and K. Fujii, Variational quantum gate optimization, arXiv preprint arXiv:1810.12745 (2018).
[45] T. Jones and S. C. Benjamin, Quantum compilation and circuit optimisation via energy dissipation, arXiv preprint arXiv:1811.03147 (2018).
[46] K. Sharma, S. Khatri, M. Cerezo, and P. J. Coles, Noise resilience of variational quantum compiling, New Journal of Physics 22, 043006 (2020).
[47] P. D. Johnson, J. Romero, J. Olson, Y. Cao, and A. Aspuru-Guzik, Qvector: an algorithm for device-tailored quantum error correction, arXiv preprint arXiv:1711.02249 (2017).
[48] X. Xu, S. C. Benjamin, and X. Yuan, Variational circuit compiler for quantum error correction, Phys. Rev. Applied 15, 034068 (2021).
[49] We note that the argument assumes a large \( \tau \) independent of the measurement accuracy. However, when the gap between the ground and the first excited state is finite, as happens in satisfactory problems, linear algebra tasks, and quantum state compiling, we only need to choose \( \tau = O(\log(1/\epsilon)) \) as a logarithmic function of the state infidelity \( \epsilon \) (when \( g(h) \) is exponential, Gaussian, or hyperbolic). Then the numerator \( N_r(O) \) and denominator \( D_r \) scales as \( O(poly(1/\epsilon)) \) and the sample complexity will not be exponentially large.
[50] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya, et al., Quantum chemistry in the age of quantum computing, Chemical reviews 119, 10856 (2019).
[51] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, Reviews of Modern Physics 92, 015003 (2020).
[52] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, Quantum algorithms for quantum chemistry and quantum materials science, Chemical Reviews 120, 12685 (2020).
[53] M. Motta and J. Rice, Emerging quantum computing algorithms for quantum chemistry (2021), arXiv:2109.02873 [quant-ph].
[54] M. A. Nielsen and I. Chuang, Quantum computation and quantum information (2002).
[55] T. Albash and D. A. Lidar, Adiabatic quantum computation, Rev. Mod. Phys. 90, 015002 (2018).
[56] L. Lin and Y. Tong, Near-optimal ground state prepara-
tion, Quantum 4, 372 (2020).
[57] X. Yuan, J. Sun, J. Liu, Q. Zhao, and Y. Zhou, Quantum simulation with hybrid tensor networks, Phys. Rev. Lett. 127, 040501 (2021).
[58] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, Hybrid quantum-classical algorithms and quantum error mitigation, Journal of the Physical Society of Japan 90, 032001 (2021), https://doi.org/10.7566/JPSJ.90.032001.
[59] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, Nature Reviews Physics 3, 625 (2021).
[60] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum (nisq) algorithms (2021), arXiv:2101.08448 [quant-ph].
[61] L. Lin and Y. Tong, Heisenberg-limited ground state energy estimation for early fault-tolerant quantum computers (2021), arXiv:2102.11340 [quant-ph].
[62] S. Endo, J. Sun, Y. Li, S. C. Benjamin, and X. Yuan, Variational quantum simulation of general processes, Phys. Rev. Lett. 125, 010501 (2020).
[63] M. Huo and Y. Li, Shallow trotter circuits fulfil error-resilient quantum simulation of imaginary time (2021), arXiv:2109.07807 [quant-ph].
[64] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nature communications 5, 1 (2014).
[65] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, New Journal of Physics 18, 023023 (2016).
[66] A. M. Childs and N. Wiebe, Hamiltonian simulation using linear combinations of unitary operations, Quantum Information and Computation 12, 10.26421/qic12.11-12 (2012).
[67] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Simulating hamiltonian dynamics with a truncated taylor series, Phys. Rev. Lett. 114, 090502 (2015).
[68] D. W. Berry, A. M. Childs, and R. Kothari, Hamiltonian simulation with nearly optimal dependence on all parameters, in 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (2015) pp. 792–809.
[69] G. H. Low and I. L. Chuang, Optimal hamiltonian simulation by quantum signal processing, Phys. Rev. Lett. 118, 010501 (2017).
[70] S. Oh, Quantum computational method of finding the ground-state energy and expectation values, Phys. Rev. A 77, 012326 (2008).
[71] T. Komà and B. Nachtergaele, The spectral gap of the ferromagnetic xxz-chain, Letters in Mathematical Physics 40, 1 (1997).
[72] R. D. Somma, Quantum eigenvalue estimation via time series analysis, New Journal of Physics 21, 123025 (2019).
[73] G. K. Karagiannidis and A. S. Lioumpas, An improved approximation for the gaussian q-function, IEEE Communications Letters 11, 644 (2007).
I. RELATED WORKS

In this section, we review the related ground-state preparation and eigenstate preparation works and compare their performances.

Variational quantum eigensolver is a major type of the ground-state preparation method [64, 65]. In the variational methods, people try to search the ground state of a Hamiltonian in a pre-determined ansatz, usually characterized by parametric quantum circuits. They then adjust and optimize the parameters in the quantum circuit based on the measured value of energy. The quantum approximate optimization algorithm [5] is another type of variational methods aiming to solve the combinatorial optimization problems based on an variational evolution combined with adiabatic evolution. The variational methods are suitable for the near-term quantum computers, since it usually requires no ancillary qubit and shallow circuits. The weakness of the variational methods is that the effectiveness always depends on the choice of the ansatz, whose validity varies for different Hamiltonian problems.

Adiabatic state preparation [55] is an experimentally-fridently non-variational ground-state method based on a time-dependent Hamiltonian evolution. In the adiabatic state preparation, people first prepare the ground state of a simple Hamiltonian \( H_0 \) and then slowly evolve it under a Hamiltonian that changes slowly from \( H_0 \) to the target Hamiltonian \( H \). Based on the adiabatic theorem, the resulting state is then close the ground state. Unlike the variational algorithms, the adiabatic algorithms are universal and valid without ansatz assumptions. When it comes to the practical usage, the adiabatic algorithm has two downsides. First, the required evolution time \( t \) depends inverse
of these methods are the number of required ancillary qubits, making it independent of the precision requirement. The query complexity the Hamiltonian energy gap and. Later on, by introducing a block-encoding method [68] and a quantum-signal-
Assumption 2 target eigenstate |ψ⟩.

deal with the original system with different time t, perform inverse Fourier transform on the ancillary qubits and then measure them. The required number of qubit ω is logarithmic to the inverse of the eigenenergy precision. The circuit depth reflected by the summation of the controlled evolution time t, is O(1/ε), where ε is the infidelity of the eigenstate preparation. The phase estimation can be improved to an iterative version with only one ancillary qubit [19]. However, similar to the adiabatic methods, the required circuit depth of the phase estimation is fundamentally O(1/ε), which is unfavorable if we want to achieve a high-precision preparation.

To further improve the efficiency of the above two nonvariational methods, the linear-combination-of-unitary (LCU) methods have been proposed [31, 32]. In Ref. [31], the authors use cos^M(H) as a spectral projector to the ground state. They decompose the Hamiltonian function cos^M(H) based on a LCU formula [66, 67], and realize the function cos^M(H) using O(log(1/Δ) log 1/ε) ancillary qubits and amplitude amplification. Here, Δ is the known lower bound of the Hamiltonian energy gap and. Later on, by introducing a block-encoding method [68] and a quantum-signal-processing technique [69] to realize a polynomial approximation of the sign function, Lin and Tong [56] further reduce the number of required ancillary qubit, making it independent of the precision requirement. The query complexity of these methods are O(1/Δ p₀ log(1/ε p₀)), which is exponentially improved comparing to the phase estimation algorithm. The difficulties of realizing these LCU or qubitization-based methods in the near-term is that, they both require many ancillary resources for LCU or block-encoding; furthermore, the realization of the select-H oracles in the LCU or qubitization approach is complicated, especially for the case of non-sparse Hamiltonian H where the number of components in H is large.

In all the above non-variational ground-state preparation methods, the efficiency relies on the following two assumptions. We remark that, the adiabatic state preparation methods also rely on a function of the state overlap p₀ which depends on the adiabatic path connecting H₀ and H.

Assumption 1 (nonvanishing state overlap assumption [31, 55, 56]). For an n-qubit gapped Hamiltonian H with the ground state |u₀⟩, we assume that it is feasible to prepare an initial state |ψ₀⟩ satisfying a nonvanishing overlap of the target eigenstate p₀ := |⟨ψ₀|u₀⟩|^2. We assume that the lower bound of the overlap satisfies p₀ ≥ Ω(Poly(1/n)).

Assumption 2 (Gapped Hamiltonian assumption [31, 55, 56]). For a n-qubit gapped Hamiltonian H with the ground state energy E₀, we assume that the energy gap E₁ − E₀ is lower bounded by a known value Δ.

Assumption 1 is reasonable in many practical scenarios, since even in many strongly-correlated quantum systems or chemical molecules Hamiltonians, the mean-field solution, like Hartree-Fock state, still has a considerable overlap with the ground state. In practice, people usually apply the variational methods [51] or adiabatic methods [70] as a heuristic way to prepare an initial state with a large overlap p₀ with the ground state. Assumption 2 is also of practical relevance for the study of quantum many-body systems. For example, the ferromagnetic XXZ Heisenberg chain H = ∑ζ σ⁺ζ σₙζ + σ⁺₀ σ⁺₁ + ∆ σ⁺₀ σ⁺₁ with J < 0 has a finite energy gap if Δ > 1 [71].

In this work, we propose a quantum algorithm to estimate the ground state energy as well as any ground-state properties with the two assumptions above. Our algorithm has the following advantages,

1. We only need one ancillary qubit in the whole algorithm. Furthermore, we do not need any quantum oracle, like select-H for LCU or qubitization methods, which could be difficult to be implemented.

2. The time complexity for the ground-state energy estimation reaches the Heisenberg limit.

3. The time complexity for the ground-state property estimation is logarithmic to the inverse of the precision requirement when we adopt the Gaussian cooling function.

To our knowledge this is the first algorithm to achieve all the advantages above. In Table II and Table III, we compare our method with several typical quantum algorithms in the two primary tasks considered in this work, i.e., energy estimation and observable estimation. Here, we mainly consider the algorithms without the usage of oracles for the block-encoding of a Hamiltonian. In the Table, we compare the maximal evolution time that characterizes the complexity of quantum circuits that realizes Hamiltonian simulation e^{−iHt}, the number of repetitions required to run the quantum circuit, and the total evolution time needed for estimating the ground state energy to within error ε. The complexity of estimating properties of excited states can be analysed in a similar fashion. Note that we simply compare with the conventional phase estimation method. Here we also use the asymptotic notations besides the usual O notation to denote the complexity up to a polylogarithmic factors, similarly to that in Ref. [31].
Here, \( N \) is called a cooling function if it satisfies,
\[ f(0) \neq 0; \]
\[ |f(h)| \leq g(h), \forall h > 0 \text{ and } \forall h' < h < 0; \]
\[ \text{lim}_{h \to 0} \frac{g(\pm h)}{g(0)} = 0, \forall h \neq 0. \]

We define the cooling operator as follows,
\[ g(\tau(E_i - E_j)) := \sum_{i=0}^{N-1} g(\tau(E_i - E_j)) |u_i\rangle \langle u_i|. \]

Then, for any given initial state \( |\psi_0\rangle \) with \( |\langle \psi_0 | u_j \rangle|^2 \neq 0 \), the cooling operator evolve \( |\psi_0\rangle \) to the eigenstate \( |u_j\rangle \),
\[ \lim_{\tau \to \infty} g(\tau(H - E_j)) |\psi_0\rangle \propto |u_j\rangle. \]

The cooling operator is usually a nonphysical operation. To realize it, we consider its dual realization based on a Fourier transform,
\[ f(x) = \int_{-\infty}^{\infty} g(h) e^{-ixh} dh, \]
\[ g(h) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{ixh} dx. \]
We define the norm of the dual function \( f(x) \) as
\[
\|f\| := \int_{-\infty}^{\infty} |f(x)| dx. \tag{11}
\]
Suppose \( \|f\| \) is finite, we can decompose \( f(x) \) to a normalized probability distribution with an extra phase,
\[
f(x) = |f(x)| e^{i\phi(x)} = \|f\| p(x) e^{i\phi(x)}. \tag{12}
\]
The Fourier transform in Eq. (10) can then be re-written as
\[
p(x) = \frac{1}{\|f\|} \int_{-\infty}^{\infty} g(h) e^{-i\phi(x)} e^{-ixh} dh,
\]
\[
g(h) = \frac{\|f\|}{2\pi} \int_{-\infty}^{\infty} p(x) e^{i\phi(x)} e^{ixh} dx. \tag{13}
\]
Based on Eq. (13), the cooling operator \( g(\tau(H - E_j)) \) in Eq. (8) can be expanded as,
\[
g(\tau(H - E_j)) = \frac{\|f\|}{2\pi} \int_{-\infty}^{\infty} p(x) e^{i\phi(x)} e^{ix\tau(H - E_j)} dx,
\]
\[
= \frac{\|f\|}{2\pi} \int_{-\infty}^{\infty} p(x) e^{i(\phi(x) - \tau x E_j)} e^{ix\tau H} dx. \tag{14}
\]
Therefore, to realize a quantum cooling operator \( g(\tau(H - E_j)) \), we can expand it to a weighted superposition of the unitary operators \( e^{i\tau H} \) with different evolution time \( t = \tau x \) ranging from \(-\infty\) to \(\infty\).

In practice, the evolution time is a limited number related to the depth of the quantum circuit. We define the truncated cooling function as,
\[
\tilde{g}(h; x_m) = \frac{\|f\|}{2\pi} \int_{-x_m}^{x_m} p(x) e^{i\phi(x)} e^{ixh} dx. \tag{15}
\]
The difference of the \( \tilde{g}(h; x_m) \) from \( g(h) \) is bounded by
\[
|g(h) - \tilde{g}(h; x_m)| = \frac{\|f\|}{2\pi} \left| \int_{-\infty}^{-x_m} p(x) e^{i\phi(x)} e^{ixh} dx + \int_{x_m}^{\infty} p(x) e^{i\phi(x)} e^{ixh} dx \right|
\]
\[
\leq \frac{\|f\|}{2\pi} \left( \int_{-\infty}^{-x_m} |p(x) e^{i\phi(x)} e^{ixh}| dx + \int_{x_m}^{\infty} |p(x) e^{i\phi(x)} e^{ixh}| dx \right)
\]
\[
= \frac{\|f\|}{2\pi} \left( 1 - \int_{-x_m}^{x_m} p(x) dx \right). \tag{16}
\]
Later in Sec. V we will show that, the estimation error caused by the finite evolution time will be bounded by the tail probability of \( p(x) \).

**Definition 2.** We say a cooling function \( g(h) \) in Definition 1 is realizable if the following requirements hold.

C0. The Fourier transform of \( g(h) \) exists,
\[
f(x) = \int_{-\infty}^{\infty} g(h) e^{-ixh} dh. \tag{17}
\]

C1. The norm of the dual function \( \|f\| \)
\[
\|f\| := \int_{-\infty}^{\infty} |f(x)| dx, \tag{18}
\]
is finite.

C2. \( 1 - \int_{-\infty}^{L(\varepsilon)} p(x) dx \leq \varepsilon, \ \forall \varepsilon \geq 0, \ \exists L(\varepsilon) = \mathcal{O}(\text{poly}(\frac{1}{\varepsilon})) \).

Now, we introduce several typical cooling functions \( g(h) \) and discuss their properties. We will focus on the cutoff cost \( L(\varepsilon) \) defined in Definition 2 which shows the hardness of realizing \( g(h) \) with real-time sampling.
A. Rectangular function

The rectangular function rect(h) is defined as,

\[ \text{rect}(h) = \begin{cases} 1, & |h| \leq \frac{1}{2}, \\ 0, & |h| > \frac{1}{2}. \end{cases} \]  \hspace{2cm} (19)

It satisfies Definition 1 and hence a cooling function. When acting as a cooling function \( g(\tau(H - E_j)) \), the rectangular function is an ideal “energy band filter”, as it projects the state to the energy subspace in the range \([E_j - \frac{1}{2\tau}, E_j + \frac{1}{2\tau}]\).

The Fourier transform of \( \text{rect}(h) \) is the Sinc function,

\[ f(x) = \text{sinc} \left( \frac{x}{2\pi} \right), \]  \hspace{2cm} (20)

where \( \text{sinc}(x) := \frac{\sin x}{x} \).

Unfortunately, the norm of \( f(x) \) is infinite,

\[ \|f\| = \int_{-\infty}^{\infty} |\text{sinc} \left( \frac{x}{2\pi} \right)| \, dx. \]

\[ = 4\pi \int_{0}^{\infty} |\text{sinc} y| \, dy \]

\[ = 4\pi \sum_{k=0}^{\infty} \int_{k\pi}^{(k+1)\pi} \frac{\sin y}{y} \, dy \]

\[ > 4\pi \sum_{k=0}^{\infty} \int_{k\pi}^{(k+1)\pi} \frac{|\sin y|}{(k+1)\pi} \, dy \]

\[ = 4\pi \sum_{k=0}^{\infty} \frac{1}{k+1}. \]  \hspace{2cm} (21)

From the divergence of \( \sum_{k=0}^{\infty} \frac{1}{k+1} \) we know that \( \|f\| = \infty \). As a result, the rectangular function is not a realizable cooling function based on Definition 2.

B. Triangular function

The triangular function tri(h) is defined as,

\[ \text{tri}(h) = \begin{cases} 1 - |h|, & |h| \leq 1, \\ 0, & |h| > 1. \end{cases} \]  \hspace{2cm} (22)

It satisfies Definition 1 and hence a cooling function. Similar to the rectangular function, it can be used as an “energy band filter”, since it filters the energy in the range \([E_j - \frac{1}{2\tau}, E_j + \frac{1}{2\tau}]\). However, it will modify the weight of different energy values in that range. When \( h \geq 0 \), the inverse function of \( \text{tri}(h) \) is

\[ \text{tri}^{-1}(p) = 1 - p \leq 1. \]  \hspace{2cm} (23)

The Fourier transform of \( \text{rect}(h) \) is the square of Sinc function,

\[ f(x) = \text{sinc}^2 \left( \frac{x}{2\pi} \right). \]  \hspace{2cm} (24)

The norm of \( f(x) \) is finite,

\[ \|f\| = 2\pi \int_{-\infty}^{\infty} \text{sinc}^2(y) \, dy = 2\pi^2. \]  \hspace{2cm} (25)

The sample probability distribution is then,

\[ p(x) = \frac{1}{\|f\|} f(x) = \frac{1}{2\pi^2} \text{sinc}^2 \left( \frac{x}{2\pi} \right). \]  \hspace{2cm} (26)
To check the realizability of cooling function \( g(h) \), we need to study the property of the tail probability of \( p(x) \) with respect to the cutoff sample time \( x_m \). Solving the following equation,

\[
\varepsilon = \int_{-x_m}^{x_m} p(x) dx,
\]

(27)

where \( p(x) \) is defined in Eq. (26), we have

\[
\varepsilon = \frac{2}{\pi} \left( \frac{\pi}{2} - \text{Si} \left( \frac{x_m}{\pi} \right) \right) + \frac{2 - 2 \cos \left( \frac{x_m}{\pi} \right)}{x_m},
\]

(28)

where \( \text{Si}(x) := \int_0^x \text{sinc}(y) dy \) is the Sinc integral function.

**Proposition 2.** We have

\[
\frac{\pi}{2} - \text{Si}(x) \leq \frac{1}{x}.
\]

(29)

**Proof.** By definition, we have,

\[
\frac{\pi}{2} - \text{Si}(x) = \int_x^\infty \frac{\sin(t)}{t} dt = \int_x^\infty \frac{\cos x \sin(t - x) + \sin x \cos(t - x)}{t} dt.
\]

(30)

Now we denote \( f(t) = \sin x \cos t + \cos x \sin t \), then

\[
\frac{\pi}{2} - \text{Si}(x) = \int_x^\infty \frac{f(t - x)}{t} dt.
\]

(31)

We introduce the Laplace transform of \( f(t) \),

\[
\mathcal{L}[f(t)](s) = \int_0^\infty f(t)e^{-st} dt,
\]

(32)

then,

\[
\frac{\pi}{2} - \text{Si}(x) = \mathcal{L} \left[ \frac{f(t - x)u(t - x)}{t} \right] (0),
\]

(33)

where \( u(t) := \eta(t \geq 0) \).

From the property of Laplace transform, we have

\[
\int_s^\infty \mathcal{L}[f(t)](p) e^{-xp} dp = \int_s^\infty \mathcal{L}[f(t - x)u(t - x)] dp = \mathcal{L} \left[ \frac{f(t - x)u(t - x)}{t} \right] (s).
\]

(34)

On the other hand, we know that

\[
\mathcal{L}[f(t)](p) = \mathcal{L}[\sin x \cos t + \cos x \sin t](p) = \frac{\sin xp + \cos x}{p^2 + 1},
\]

(35)

then

\[
\int_s^\infty \mathcal{L}[f(t)](p) e^{-xp} dp = \int_s^\infty \frac{\sin xp + \cos x}{p^2 + 1} e^{-xp} dp.
\]

(36)

Taking \( s = 0 \) and combine Eq. (33), (34), and (36), we have

\[
\frac{\pi}{2} - \text{Si}(x) = \int_0^\infty \frac{\sin xp + \cos x}{p^2 + 1} e^{-xp} dp.
\]

(37)

Using Cauchy-Schwarz inequality we can prove that,

\[
p \sin x + 1 \cdot \cos x \leq \sqrt{1 + p^2} \sqrt{\sin^2 x + \cos^2 x} = \sqrt{1 + p^2},
\]

(38)
\[
\frac{\pi}{2} - \text{Si}(x) = \int_0^\infty \frac{\sin xp + \cos x}{p^2 + 1} e^{-xp} dp
\]
\[
\leq \int_0^\infty \frac{1}{\sqrt{p^2 + 1}} e^{-xp} dp
\]
\[
\leq \int_0^\infty e^{-xp} dp
\]
\[
= \frac{1}{x}.
\]

Therefore,
\[
\varepsilon \leq \frac{2}{\pi} \frac{\pi}{x_m} + \frac{2 - 2 \cos \left( \frac{x_m x}{\pi} \right)}{x_m}
\]
\[
\leq \frac{2}{x_m} + \frac{4}{x_m} = \frac{6}{x_m}.
\]

We have
\[
x_m \leq \frac{6}{\varepsilon} =: L(\varepsilon) = \mathcal{O}(\text{poly}(\frac{1}{\varepsilon})�).
\]

Therefore, the triangle function is a realizable cooling function. However, the circuit depth requirement of triangle function is still demanding, as the tail probability \( \varepsilon \) decays slowly with the cutoff time \( x_m \).

C. Double-side exponential decay function

The double-side exponential decay function is
\[
g(h) := e^{-|h|}.
\]

Hereafter we call it exponential function for simplicity. It satisfies Definition 1 and hence a cooling function. The exponential function is of special importance as it describes the imaginary time evolution \( e^{-\tau H} \) when the eigenenergies are all positive. When \( h \geq 0 \), the inverse function is
\[
g^{-1}(p) = \ln \left( \frac{1}{p} \right).
\]

The Fourier transform of the exponential function is
\[
f(x) = \frac{2}{1 + x^2}.
\]

The norm of \( f(x) \) is finite,
\[
\|f\| = \int_{-\infty}^{\infty} \frac{2}{1 + x^2} dx = 2\pi.
\]

The sample probability distribution is,
\[
p(x) = \frac{1}{\|f\|} f(x) = \frac{1}{2\pi} \frac{2}{1 + x^2},
\]
which is a Lorentian distribution.

Now, we check the tail probability of \( p(x) \) with respect to the cutoff sample time \( x_m \). Solving the following equation,
\[
\varepsilon = \int_{-x_m}^{x_m} p(x) dx,
\]
where \( p(x) \) is defined in Eq. (46), we have

\[
x_m = \tan \left[ \frac{\pi}{2} (1 - \varepsilon) \right],
\]

(48)

To calculate the dependence of \( x_m \) with respect to \( \frac{1}{\varepsilon} \), we expand \( \frac{1}{x_m} \) at the point \( \varepsilon \to 0^+ \) using Taylor expansion with reminder,

\[
\frac{1}{x_m} = \frac{1}{\tan \left[ \frac{\pi}{2} (1 - \varepsilon) \right]} = 0 + \frac{\pi}{2} \left( \frac{1}{\cos^2 \left( \frac{\pi}{2} \cdot 0 \right)} \right) \varepsilon + \frac{\pi^2}{4} \frac{\sin(\pi \xi)}{\cos^4 \left( \frac{\pi}{2} \xi \right)} \frac{\varepsilon^2}{2}
\]

\[
> \frac{\pi}{2} \varepsilon,
\]

(49)

where \( \xi \in (0, \varepsilon) \). We then have,

\[
x_m < \frac{2}{\pi} \frac{1}{\varepsilon} =: L(\varepsilon) = \mathcal{O}(\text{poly}(\frac{1}{\varepsilon})).
\]

(50)

Therefore, the exponential function is a realizable cooling function. Still, the circuit depth requirement is demanding, as the tail probability \( \varepsilon \) decays slowly with the cutoff time \( x_m \). In what follows, we consider two cooling functions with small tails.

D. Gaussian function

The Gaussian function we used here is

\[
g(h) := e^{-h^2}.
\]

(51)

It satisfies Definition 1 and hence a cooling function. When \( h \geq 0 \), the inverse function is

\[
g^{-1}(p) = \sqrt{\ln \left( \frac{1}{p} \right)}.
\]

(52)

The Fourier transform of Gaussian function is still a Gaussian function,

\[
f(x) = \sqrt{\pi} e^{-\frac{x^2}{4}}.
\]

(53)

The norm of \( f(x) \) is finite,

\[
\|f\| = \sqrt{\pi} \int_{-\infty}^{\infty} e^{-\frac{x^2}{4}} dx = 2\pi.
\]

(54)

The sample probability distribution is,

\[
p(x) = \frac{1}{\|f\|} f(x) = \frac{1}{2\pi} \sqrt{\pi} e^{-\frac{x^2}{4}},
\]

(55)

which is a Gaussian distribution.

Now, we check the tail probability of \( p(x) \) with respect to the cutoff sample time \( x_m \). Solving the following equation,

\[
\varepsilon = \int_{-x_m}^{x_m} p(x) dx,
\]

(56)

where \( p(x) \) is defined in Eq. (55), we have

\[
\varepsilon = \text{erfc}(\frac{x_m}{2}),
\]

(57)
where \( \text{erfc}(x) := 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \) is the complementary of the error function. It can be upper-bounded by a Chernoff-type formula \([73]\),

\[
\text{erfc}(x) \leq \frac{1}{\sqrt{\pi x}} e^{-x^2}, \quad x \geq 0.
\]

Therefore,

\[
\varepsilon = \text{erfc}\left(\frac{x_m}{2}\right) \leq \frac{2}{\sqrt{\pi x_m}} e^{-\frac{x_m^2}{4}}
\]
\[
\Rightarrow \ln \frac{\sqrt{\pi} \varepsilon}{2} \leq -\frac{x_m^2}{4} - \ln x_m
\]
\[
\Rightarrow \frac{x_m^2}{4} \leq \frac{x_m^2}{4} + \ln x_m \leq -\ln \left(\frac{\sqrt{\pi} \varepsilon}{2}\right)
\]
\[
\Rightarrow x_m \leq 2\sqrt{-\ln \varepsilon - \ln \frac{\sqrt{\pi}}{2}} \leq 2\sqrt{\ln \left(\frac{1}{\varepsilon}\right)}.
\]

We have \( L(\varepsilon) = 2\sqrt{\ln \left(\frac{1}{\varepsilon}\right)} = \mathcal{O}(\text{poly}(\frac{1}{\varepsilon})). \) Therefore, the Gaussian function is a realizable cooling function. Note that, unlike the triangular function or the exponential function, the tail of Gaussian function decays quickly with respect to \( x_m \). This implies that the Gaussian cooling is more experimentally friendly.

**E. Secant hyperbolic function**

The secant hyperbolic function is

\[
\text{sech}(h) := \frac{2}{e^h + e^{-h}}.
\]

It satisfies Definition 1 and hence a cooling function. The inverse of \( \text{sech} \) is hard to be presented in an analytical form, but we can easily derive an upper bound for it,

\[
g^{-1}(p) < -\ln \frac{2}{p}.
\]

The Fourier transform of secant hyperbolic function is still a secant hyperbolic function,

\[
f(x) = \pi \text{sech} \left(\frac{\pi}{2} x\right).
\]

The norm of \( f(x) \) is finite,

\[
\|f\| = \pi \int_{-\infty}^{\infty} \text{sech} \left(\frac{\pi}{2} x\right) dx = 2\pi.
\]

The sample probability distribution is,

\[
p(x) = \frac{1}{\|f\|} f(x) = \frac{1}{2\pi} \pi \text{sech} \left(\frac{\pi}{2} x\right),
\]

which is a secant hyperbolic distribution.

Now, we check the tail probability of \( p(x) \) with respect to the cutoff sample time \( x_m \). Solving the following equation,

\[
\varepsilon = \int_{-x_m}^{x_m} p(x) dx,
\]

where \( p(x) \) is defined in Eq. (64), we have

\[
x_m = \frac{2}{\pi} \ln \left\{\tan \left[\frac{\pi}{2} \left(1 - \frac{\varepsilon}{2}\right)\right]\right\}
\]
\[
\Rightarrow e^{-\frac{\pi}{2} x_m} = \frac{1}{\tan \left[\frac{\varepsilon}{2} \left(1 - \frac{\varepsilon}{2}\right)\right]}.
\]
Similar to Eq. (49), we expand the inversed tangent function at the point \( \varepsilon \to 0^+ \) using Taylor expansion with reminder,

\[
e^{-\frac{\pi}{2}x_m} = \tan \left[ \frac{\pi}{2} (1 - \frac{\xi}{2}) \right] = 0 + \frac{\pi}{4} \frac{1}{\cos^2 \left( \frac{\pi}{4} \cdot 0 \right)} \varepsilon + \pi^2 \frac{\sin^4 \left( \frac{\pi}{2} \xi \right)}{\sin^3 \left( \frac{\pi}{2} \xi \right)} \frac{\varepsilon^2}{2} > \frac{\pi}{4} \varepsilon,
\]

where \( \xi \in (0, \varepsilon) \). Therefore,

\[
x_m < \frac{2}{\pi} \ln \left( \frac{4}{\pi \varepsilon} \right) =: L(\varepsilon) = \mathcal{O}(\text{poly}(\frac{1}{\varepsilon})). \tag{68}
\]

Therefore, the secant hyperbolic function is also a realizable function. The overhead of secant hyperbolic function is better than the Gaussian function; however, the asymptotic cost of the Gaussian function is quadratically better than the secant hyperbolic function.

### III. DETAILED QUANTUM COOLING ALGORITHM

Suppose we want to estimate the expectation value on an eigenstate \(|u_j\rangle, \langle O \rangle := \langle u_j | O | u_j \rangle\), using the quantum cooling algorithm. For a given finite evolution time \( \tau \), the estimation value of \( \langle O \rangle_\tau \) is given by

\[
\langle O \rangle_\tau = \frac{N_\tau(O)}{D_\tau}, \tag{69}
\]

where

\[
D_\tau = \langle \psi_0 | g^2(\tau(H - E_j)) | \psi_0 \rangle,
N_\tau(O) = \langle \psi_0 | g(\tau(H - E_j)) Og(\tau(H - E_j)) | \psi_0 \rangle. \tag{70}
\]

are, respectively, the normalization factor and the unnormalized expectation value. To apply the Hadamard test, we decompose the observable \( O \) by the Pauli operators,

\[
O = \sum_{l \in P_n} o_l P_l = \|O\|_1 \sum_{l \in P_n} \Pr_O(l) P_l \tag{71}
\]

where \( P_n \) denotes the \( n \)-qubit Pauli group. \( o_l \) is the coefficient of the Pauli component \( P_l \). We remark that, the coefficients \( \{o_l\} \) are all set to be positive. The signs of the coefficients are put into the corresponding Pauli matrices \( \{P_l\} \). In Eq. (71), \( \|O\|_1 \) is the \( l_1 \)-norm of the Pauli coefficients of \( O \),

\[
\|O\|_1 = \sum_{l \in P_n} o_l. \tag{72}
\]

The probability distribution \( \Pr_O(l) \) is defined to be,

\[
\Pr_O(l) = \frac{o_l}{\sum_{l \in P_n} o_l} = \frac{1}{\|O\|_1} o_l. \tag{73}
\]

Using Eq. (14), we expand the cooling operator \( g(\tau(H - E_j)) \) and receive the following estimation formulas,

\[
D_\tau = \left( \frac{\|f\|}{2\pi} \right)^2 \int_{-\infty}^{\infty} dy \tilde{p}(y) e^{-iyE_j} \langle \psi_0 | e^{iyH} \psi_0 \rangle,
N_\tau(O) = \|O\|_1 \left( \frac{\|f\|}{2\pi} \right)^2 \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' p(x)p(x') \sum_{l \in P_n} \Pr_O(l) e^{-i\tau(x-x')} E_j \langle \psi_0 | e^{-i\tau x' H} P_l e^{i\tau x' H} \psi_0 \rangle, \tag{74}
\]

where \( \tilde{p}(y) = \int_{-\infty}^{\infty} dx p(x) e^{iyx} \).
where \( \hat{p}(y) := \frac{1}{2} \int_{-\infty}^{\infty} p(\frac{z+y}{2})p(\frac{z-y}{2})dz \) is the self-correlation of the probability \( p(x) \). We remark that, the normalization factor \( \left( \frac{\|f\|_2}{2\pi} \right)^2 \) can be removed, since it is the same for both \( D_\tau \) and \( N_\tau(O) \), hence independent of the estimation of \( \langle O \rangle \). In the following discussion, we ignore this normalization factor during the estimation procedure.

In the quantum cooling algorithm, we generate the (normalized) evolution time \( x \) (or \( y \)) based on the sample probability \( p(x) \) (or \( \hat{p}(y) \)) and sample the inner product values using quantum experiments. For a practical consideration, when the (normalized) evolution time \( x \) (or \( y \)) is larger than a cutoff value \( x_m \), we don’t perform the quantum experiment and direct denote the estimation of this round to be 0. In this way, the estimation formula of \( N_\tau(O) \) and \( D_\tau \), originally given by Eq. (74), now becomes

\[
\begin{align*}
D_\tau(x_m) &= \int_{-x_m}^{x_m} dy \hat{p}(y)e^{-i\tau y}E_j(\langle \psi_0|e^{i\tau yH}\psi_0 \rangle), \\
N_\tau(x_m)(O) &= \|O\|_1 \int_{-x_m}^{x_m} dx \int_{-x_m}^{x_m} dx' p(x)p(x') \sum_{l \in P_n} Pr_O(l)e^{-i\tau x' H}P_l e^{i\tau x' H}\langle \psi_0|e^{-i\tau x' H}P_l e^{i\tau x'H}\psi_0 \rangle.
\end{align*}
\tag{75}
\]

To estimate the values of \( N_\tau(x_m)(O) \) and \( D_\tau(x_m) \) in Eq. (75), the core issue is to realize the unbiased estimation of the following quantities,

\[
\begin{align*}
\langle \psi_0|e^{i\tau y H}\psi_0 \rangle, \\
\langle \psi_0|e^{-i\tau x H}P_l e^{i\tau x'H}\psi_0 \rangle.
\end{align*}
\tag{76}
\]

These \( \langle \psi|U|\psi \rangle \) form quantities can be estimated using the Hadamard test, shown in Fig. 3. Here, the \( S \) gate is

\[
S = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix}.
\tag{77}
\]

To measure \( \langle \psi|U|\psi \rangle \), we first prepare the state \( |\psi \rangle \) and an extra ancillary qubit prepared on \( |+ \rangle \). Afterward, we perform a \( C-U \) gate from ancillary to \( |\psi \rangle \). If we directly measure the ancillary qubit on the \( X \)-basis, the outcome 0 will be 0 with a probability of \( \frac{1}{2}(1 + \text{Re}(\langle \psi|U|\psi \rangle)) \) and 1 with a probability of \( \frac{1}{2}(1 - \text{Re}(\langle \psi|U|\psi \rangle)) \). Alternatively, if we perform an extra \( W \) gate before the \( X \)-basis measurement, the outcome 0 will be 0 with a probability of \( \frac{1}{2}(1 + \text{Im}(\langle \psi|U|\psi \rangle)) \) and 1 with a probability of \( \frac{1}{2}(1 - \text{Im}(\langle \psi|U|\psi \rangle)) \).

![FIG. 3. The diagram of the Hadamard test. If \( b = 0 \), \( \text{Pr}(a = 0) = \frac{1}{2}(1 + \text{Re}(\langle \psi|U|\psi \rangle)) \); if \( b = 1 \), \( \text{Pr}(a = 0) = \frac{1}{2}(1 + \text{Im}(\langle \psi|U|\psi \rangle)). \)](image)

To simplify the theoretical analysis, we now introduce a combined estimation of the real part and imaginary part in a single round. In each round of experiment, we first randomly decide the binary value \( b \) uniformly from \( \{0, 1\} \) in Fig. 3. We denote the binary uniform distribution of \( b \) as \( \text{Pr}_U(b) \). Based on the measurement outcome \( a \), we construct the following complex-valued estimator,

\[
\hat{r} = \begin{cases} 2, & (b, a) = (0, 0), \\
-2, & (b, a) = (0, 1), \\
2i, & (b, a) = (1, 0), \\
-2i, & (b, a) = (1, 1). \\
\end{cases}
\tag{78}
\]
That is, $\hat{r} = 2^b(-1)^a$. In this case, we have

$$
\mathbb{E}(\hat{r}) = \sum_{b \in \{0,1\}} \Pr(U) \Pr(a|b) \hat{r}(b,a) \\
= \frac{1}{4} \left(1 + \text{Re}(\langle \psi|U|\psi\rangle)\right) \cdot 2 + \frac{1}{4} \left(1 - \text{Re}(\langle \psi|U|\psi\rangle)\right) \cdot (-2) \\
+ \frac{1}{4} \left(1 + \text{Im}(\langle \psi|U|\psi\rangle)\right) \cdot (2i) + \frac{1}{4} \left(1 - \text{Im}(\langle \psi|U|\psi\rangle)\right) \cdot (-2i) \\
= \text{Re}(\langle \psi|U|\psi\rangle) + i\text{Im}(\langle \psi|U|\psi\rangle) \\
= \langle \psi|U|\psi\rangle.
$$

Therefore, $\hat{r}$ is an unbiased estimator for $\langle \psi|U|\psi\rangle$.

In the estimation of $D_{\tau}$ and $N_{\tau}(O)$, we also put the phase term and normalization factor $\|O\|_1$ in Eq. (74) into the estimator. That is, we define

$$
\hat{d} = e^{-i\tau y E_j} \hat{r}, \\
\hat{n} = \|O\|_1 e^{-i\tau (x-x') E_j} \hat{r}.
$$

As a result, if we randomly sample $y$ from $\tilde{p}(y)$, based on Eq. (74) we have

$$
\mathbb{E}_{y,b,a}(\hat{d}) = D_{\tau}.
$$

On the other hand, if we randomly sample $x,x'$ from $p(x)$ and $l$ from $\text{Pr}_O(l)$, based on Eq. (74) we have

$$
\mathbb{E}_{x,x',l,b,a}(\hat{n}) = N_{\tau}(O).
$$

Algorithm 1 below is based on constructing the estimator in Eq. (83). We also introduce the truncation into the sampling of $y$.

In the main text, we consider two applications of the universal cooling: 1) to estimate the ground state properties when the ground state energy $E_0$ is known; 2) to estimate the properties of the $j$-th eigenstate without the exact knowledge of the energy. We can divide it into three elementary tasks:

1. Estimate the eigenenergy $E_j$ of the $j$-th eigenstate given a initial guess interval $[E_j^L, E_j^U]$;
2. Given a known eigenenergy $E_j$, estimate the normalization factor $D_{\tau}$;
3. Given a known eigenenergy $E_j$ and an observable $O$, estimate the normalization factor $N_{\tau}(O)$.

We now introduce the detailed algorithm for these tasks.

### A. Estimate the normalization factor

We first introduce the algorithm to estimate the normalization factor $D_{\tau}$ based on Eq. (75) and the Hadamard test, when the eigenenergy $E_j$ is known. Recall that the unbiased estimator for the normalization factor is,

$$
\hat{d} = e^{-i\tau y E_j} \hat{r} = 2^b(-1)^a e^{-i\tau y E_j},
$$

If we randomly sample $y$ from $\tilde{p}(y)$, based on Eq. (74) we have

$$
\mathbb{E}_{y,b,a}(\hat{d}) = D_{\tau}.
$$

Algorithm 1 below is based on constructing the estimator in Eq. (83). We also introduce the truncation into the sampling of $y$. 
Algorithm 1 Normalization factor estimation

Input: An n-qubit Hamiltonian $H$; initial state $|\psi_0\rangle$ with nonzero overlap with $j$-th eigenstate of $H$: $p_j = |\langle u_j | \psi_0 \rangle|^2$; the energy $E_j$ for the $j$-th eigenstate; cooling function $g(h)$ and the corresponding sampling probability $p(x)$; proper choice of the imaginary time $\tau$, normalized cutoff time $x_m$.

Output: Estimation of the normalization factor $D_{\tau}$.

1: for $p = 1$ to $N_M$ do
2: Sample a normalized evolution time $y_p$ from $\bar{p}(y) := \frac{1}{\tau} \int_{-\infty}^{\infty} p\left(\frac{2 \pi y}{\tau} \right) p\left(\frac{2 \pi y}{\tau}\right) dy$.
3: if $y_p > x_m$ then
   \hspace{1cm} $\triangleright$ Exceed the preset cutoff value
   4: Set the estimation $\hat{d}_p = 0$.
5: else
   \hspace{1cm} $\triangleright$ Normal single-shot quantum sampling by Hadamard test
6: Prepare an ancillary qubit on $|+\rangle$ state and the initial state $|\psi_0\rangle$.
7: Implement a controlled-unitary $C-U$ from the ancillary qubit to the state $|\psi_0\rangle$. Here, $U = e^{i\tau y_p H}$.
8: Generate a random bit $b_p$ with a uniform distribution of values $\{0, 1\}$. Perform a gate $W_{b_p}$ on the ancillary qubit.
9: Measure the ancillary qubit on $X$-basis, and record the binary result $a_p$. Then record the values $\{y_p, b_p, a_p\}$ and set the estimation value $\hat{d}_p := 2(-1)^{a_p} e^{i\tau y_p E_j}$.
10: end if
11: end for
12: Calculate the estimated normalization factor $\hat{D}_{\tau}^{(x_m)} := \text{Re} \left( \frac{1}{N_M} \sum_{p=1}^{N_M} \hat{d}_p \right)$.

B. Estimate the eigenenergy

Suppose the initial state $|\psi_0\rangle$ has a constant overlap with the eigenstate $|u_j\rangle$. Then we can sweep the parameter $E_j^{(e)}$ in a range to maximize $D_{\tau}$. The estimation accuracy depends on the finite imaginary time $\tau$, finite truncation $x_m$, sampling number $N_M$ and the overlap $p_j$.

To clarify this, we consider a simple case where the $j$-th eigenenergy is known to be in a range $E_j \in [E^L_j, E^U_j]$; moreover, we suppose other eigenenergies is far from this range. We expand the initial state in the eigenstate basis,

$$|\psi_0\rangle = \sum_{j=0}^{N-1} c_j |u_j\rangle. \quad (85)$$

We denote the square overlap of $|\psi\rangle$ and $|u_i\rangle$ to be

$$p_i = |\langle u_i | \psi \rangle|^2 = |c_i|^2. \quad (86)$$

For an energy value $E$ in the range $[E^L_j, E^U_j]$, we calculate the ideal normalization factor $D_{\tau}$,

$$D_{\tau}(E) = |\langle \psi | g^2(\tau(H - E)) \psi \rangle|$$
$$= g^2(\tau(E_j - E)) p_j + \sum_{i \neq j} g^2(\tau(E_i - E)) p_i$$
$$\approx g^2(\tau(E_j - E)) p_j. \quad (87)$$

The approximation holds when $g^2(\tau(E_j - E)) p_j \gg g^2(\tau(E_i - E)) p_i$. This naturally holds when the eigenenergies $\{E_i\}_{i \neq j}$ is far from the range $[E^L_j, E^U_j]$. From Eq. (87) we can see that, the normalization factor takes the local maximum value close to $p_j$ when the energy value $E = E_j$,

$$E_j = \arg\max_{E \in [E^L_j, E^U_j]} D_{\tau}(E). \quad (88)$$

Therefore, we can sweep the values of $E$ in a range $[E^L_j, E^U_j]$ to search the largest value of $D_{\tau}(E)$. The eigenenergy searching algorithm is in Algorithm 2. In the algorithm, we assume a free usage of computational resources. This can be improved by introducing better peak-value searching algorithm. We leave the improvement of the classical search process for future works.
Estimation of the eigenenergy

If we randomly sample an eigenstate property estimation is shown in Algorithm 3.

Recall the unbiased estimator of \( N_x \) is

\[
\hat{n} = \|O\|_1 e^{-i\tau(x-x')}E_j f = 2\|O\|_1 e^{-i\tau(x-x')}E_j E_l b (-1)^a.
\]

If we randomly sample \( x, x' \) from \( p(x) \) and \( l \) from \( Pr_O(l) \), based on Eq. (74) we have

\[
E_{x,x',l,b,a}(\hat{n}) = N_x(O).
\]

We now introduce the estimation algorithm of \( N_x(O) \) based on Eq. (75) and the Hadamard test. The unnormalized eigenstate property estimation is shown in Algorithm 3.

To summarize, we list the algorithm for the two applications in the main text, shown in Algorithms 4 and 5.
Algorithm 3 Unnormalized eigenstate property estimation

Input: An n-qubit Hamiltonian \( H \); initial state \( |\psi_0\rangle \) with nonzero overlap with \( j \)-th eigenstate of \( H \): \( p_j = |\langle u_j|\psi_0\rangle|^2 \); the energy \( E_j \) for the \( j \)-th eigenstate; the target observable \( \hat{O} = \sum_{l \in \mathbb{R}} \alpha_l P_l \); cooling function \( g(h) \) and the corresponding sampling probability \( p(x) \); proper choice of the imaginary time \( \tau \); normalized cutoff time \( x_m \).

Output: Estimation of the unnormalized observable value \( N_t(O) \).

1: for \( q = 1 \) to \( N_M \) do
2: Sample two independent variables \( x_q \) and \( x'_q \) from \( p(x) \).
3: if \( x_q > x_m \) or \( x'_q > x_m \) then \( \triangleright \) Exceed the preset cutoff value
4: Set the estimation \( \hat{n}_q = 0 \).
5: else \( \triangleright \) Normal single-shot quantum sampling by Hadamard test
6: Randomly sample a Pauli string \( P_l \) from the support of \( O \), based on the probability distribution \( \Pr_{O}(l) = \alpha_l / \| O \|_1 \).
7: Prepare an ancillary qubit on \( |+\rangle \) state and the initial state \(|\psi_0\rangle \).
8: Implement a controlled-unitary \( C-V \) from the ancillary qubit to the state \(|\psi_0\rangle \). Here, \( V = e^{i\tau x_q H} P e^{i\tau x_q H} \).
9: Generate a random bit \( b_q \) with a uniform distribution of values \{0, 1\}. Perform a gate \( W^{b_q} \) on the ancillary qubit.
10: Measure the ancillary qubit on X-basis, and record the binary result \( b_q \). Then set the estimation value \( \hat{n}_q = 2\|O\|_1 (1-1)^{b_q} e^{-i\tau(x_q-x'_q)} H \).
11: \end if
12: \end for
13: Set the estimation of the unnormalized expectation value \( \hat{N}_\psi^{(x_m)}(O) := \text{Re} \left( \frac{1}{N_M} \sum_{q=1}^{N_M} \hat{n}_q \right) \).

Algorithm 4 Ground state property estimation (with known ground-state energy)

Input: An n-qubit Hamiltonian \( H \); initial state \( |\psi_0\rangle \) with nonzero overlap with the ground state of \( H \): \( p_0 = |\langle u_0|\psi_0\rangle|^2 \); the ground state energy \( E_0 \); the target observable \( \hat{O} = \sum_{l \in \mathbb{R}} \alpha_l P_l \); cooling function \( g(h) \) and the corresponding sampling probability \( p(x) \); proper choice of the imaginary time \( \tau \); normalized cutoff time \( x_m \).

Output: Estimation of the observable value \( \langle u_0|O|u_0\rangle \) for the ground state.

1: Perform Algorithm 1 to output the estimation of the normalization factor \( \hat{D}^{(x_m)} \).
2: Perform Algorithm 3 to output the estimation of the unnormalized observable value \( \hat{N}_\psi^{(x_m)} \).
3: Output the estimation \( \langle \hat{O} \rangle_{\psi(\tau)} = \hat{N}_\psi^{(x_m)}(O) / \hat{D}_\psi^{(x_m)} \).

IV. ERROR AND RESOURCE REQUIREMENT ANALYSIS FOR THE EIGENENERGY ESTIMATION

In this section, we study the estimation error of the \( j \)-th eigenenergy using Algorithm 2. Based on the error dependence, we estimate the resource requirements (i.e., circuit depth and sample number) of the energy estimation. For simplicity, we focus on the case when the following assumptions holds,

1. The cooling function \( g(h) \) is even and real. Therefore,
   \( g(\tau(H-E)) = g(\tau(H-E)) \).

(94)

Algorithm 5 Eigenstate property estimation (with unknown eigenstate energy)

Input: An n-qubit Hamiltonian \( H \); initial state \( |\psi_0\rangle \) with nonzero overlap with the ground state of \( H \): \( p_0 = |\langle u_0|\psi_0\rangle|^2 \); the energy interval \( [E_j^L, E_j^U] \) for the \( j \)-th eigenstate; the target observable \( \hat{O} = \sum_{l \in \mathbb{R}} \alpha_l P_l \); cooling function \( g(h) \) and the corresponding sampling probability \( p(x) \); proper choice of the imaginary time \( \tau \); normalized cutoff time \( x_m \).

Output: Estimation of the observable value \( \langle u_j|O|u_j\rangle \) for the \( j \)-th eigenstate.

1: Perform Algorithm 2 to output the estimation of the eigenenergy \( E_j \) and the normalization factor \( \hat{D}^{(x_m)} \).
2: Perform Algorithm 3 to output the estimation of the unnormalized observable value \( \hat{N}_\psi^{(x_m)} \).
3: Output the estimation \( \langle \hat{O} \rangle_{\psi(\tau)} = \hat{N}_\psi^{(x_m)}(O) / \hat{D}_\psi^{(x_m)} \).
2. There is only one eigenenergy in the range $[E^L_j, E^U_j]$. The value of $g(\tau(E - E_i))^2$ when $E_i \neq E_j$ is negligible.

Under these assumptions, the normalization factor can then be expressed as,

$$D_{\tau}(E) = \langle \psi_0 | g(\tau(H - E))^2 | \psi_0 \rangle$$

$$= \sum_{j=0}^{N-1} g(\tau(E_j - E))^2 | \langle u_j | \psi_0 \rangle |^2$$

$$= \sum_{j=0}^{N-1} g(\tau(E_j - E))^2 p_j$$

$$\approx p_j g(\tau(E - E_j))^2.$$ 

(95)

Here, the approximation holds when the values $g(\tau(E - E_i))^2$ contributed by other eigenenergies $E_i$ are negligible. In the following Proposition 3, we will make this approximation rigorous. The location of the peak of $D_{\tau}(E)$ then indicates the eigenenergy $E_j$.

In practice, however, we can only obtain the estimation $\hat{D}_{\tau}(x_m) (E)$ of $D_{\tau}(E)$, considering the finite cutoff time $x_m$ and finite sample number $N_M$. We are going to prove that, the solution of the following maximization problem

$$\hat{E}_j := \arg \max_{E \in [E^L_j, E^U_j]} \hat{D}_{\tau}(x_m) (E),$$

(96)

will be close to the real solution $E_j$.

In Fig. 4 we summarize our analysis. In Sec. IV A, we bound the distance between $D_{\tau}(E)$ and $p_j g(\tau(E - E_j))^2$ caused by finite $\tau$ when $E \in [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}]$. In Sec. IV B, we bound the estimation error of the normalization factor $D_{\tau}(E)$ caused by the normalized cutoff time $x_m$. In Sec. IV C, based on the measurement using Hadamard test, we bound the statistical error of the estimation of $D_{\tau}(x_m)(E)$ caused by the finite sampling error $N_M$. Finally, in Sec. IV D, we consider the peak-value search problem using the estimator and analyze the eigenenergy accuracy dependence to the circuit depth and sample complexity. We will see how the cooling bandwidth $g^{-1}(1-\varepsilon)$ affects the accuracy of the peak-value search.

A. Finite imaginary evolution time

We first bound the distance between $D_{\tau}(E)$ and $p_j g(\tau(E - E_j))^2$ caused by finite $\tau$ when $E \in [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}]$. 

FIG. 4. Summary of the error analysis of eigenenergy estimation. We first study the effect of the finite imaginary time $\tau$, finite normalized cutoff time $x_m$ and finite sampling number $N_M$ on the normalization estimation, and then bound the difference between the estimation value $\hat{E}_j$ from true value $E_j$. 

**Proposition 3** (Error of the normalization function introduced by finite imaginary time). When \( E \in [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}] \), we have

\[
|D_r(E) - p_j g(\tau(E - E_j))|^2 \leq \varepsilon_r^1,
\]

(97)

when \( \tau \geq \frac{2}{\Delta} g^{-1}(\frac{\varepsilon_r^1}{2}) \). When \( E = E_j \), we can improve the requirement to \( \tau \geq \frac{1}{\Delta} g^{-1}(\frac{\varepsilon_r^1}{2}) \).

**Proof.** Denote the projector \( \hat{P}_j = |u_j\rangle \langle u_j| \). We first bound the distance between \( g(\tau(H - E)) \) and \( g(\tau(E_j - E))\hat{P}_j \). Note that,

\[
g(\tau(H - E)) - g(\tau(E_j - E))\hat{P}_j = \sum_{i \neq j} g(\tau(E_i - E)) |u_i\rangle \langle u_i|.
\]

(98)

When \( E \in [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}] \), we have \( |E - E_j| \geq \frac{\Delta}{2} \). In this case, when \( \tau \geq \frac{2}{\Delta} g^{-1}(\frac{\varepsilon_r^1}{2}) \), the coefficient \( g(\tau(E_i - E)) \) is bounded by

\[
g(\tau(E_i - E)) = g\left(\frac{2|E_i - E|}{\Delta} g^{-1}(\frac{\varepsilon_r^1}{2})\right) \\
\leq g\left(g^{-1}(\frac{\varepsilon_r^1}{2})\right) \\
= \frac{\varepsilon_r^1}{2}.
\]

(99)

In the second line, we use the fact that \( g(h') \leq g(h) \) when \( h' > h > 0 \) in Definition 1. When \( E = E_j \), we have \( |E_j - E_i| \geq \Delta \). In this case, we can achieve the same precision in Eq. (99) using \( \tau \geq \frac{1}{\Delta} g^{-1}(\frac{\varepsilon_r^1}{2}) \).

Combining Eq. (98) and Eq. (99), we have

\[
\|g(\tau(H - E)) - g(\tau(E_j - E))\hat{P}_j\|_\infty \leq \frac{\varepsilon_r^1}{2}.
\]

(100)

When \( \tau \geq \frac{2}{\Delta} g^{-1}(\frac{\varepsilon_r^1}{2}) \). Here, \( \|A\|_\infty \) denotes the spectral norm of matrix \( A \).

Now, we bound the distance between \( D_r(E) \) and \( p_j g(\tau(E - E_j))^2 \),

\[
|D_r(E) - p_j g(\tau(E - E_j))|^2 = |\langle \psi_0| g(\tau(H - E))g(\tau(H - E))|\psi_0\rangle - g(\tau(E - E_j))^2 \langle \psi_0|\hat{P}_j\hat{P}_j|\psi_0\rangle | \\
\leq |\langle \psi_0| g(\tau(H - E))g(\tau(H - E))|\psi_0\rangle - g(\tau(E - E_j)) \langle \psi_0|g(\tau(H - E))\hat{P}_j|\psi_0\rangle | \\
+ g(\tau(E - E_j)) |\langle \psi_0|g(\tau(H - E))\hat{P}_j|\psi_0\rangle - g(\tau(E - E_j)) \langle \psi_0|\hat{P}_j\hat{P}_j|\psi_0\rangle | \\
\leq \frac{\varepsilon_r^1}{2} + g(\tau(E - E_j)) \frac{\varepsilon_r^1}{2} \leq \varepsilon_r^1.
\]

(101)

\[\square\]

**B. Finite normalized cutoff time**

We then bound the estimation error of \( D_r(E) \) caused by the normalized cutoff time \( x_m \). The estimation formula is given in Eq. (75),

\[
D_r^{(x_m)}(E) = \int_{-x_m}^{x_m} dy \bar{p}(y)e^{-i\gamma y E_j} \langle \psi_0|e^{i\gamma y H}|\psi_0\rangle.
\]

(102)

Here, \( \bar{p}(y) = \int_{-\infty}^{\infty} p(t - y)p(t)dt = [p * p](y) \).
Proposition 4 (Error of the normalization factor introduced by finite normalized cutoff time). The truncated estimation \( D_r^{(x_m)}(E) \) defined in Eq. (102) is related to the normalization factor defined in Eq. (74) by

\[
|D_r^{(x_m)}(E) - D_r(E)| \leq \varepsilon_x^{(1)} , \forall E \in \mathbb{R},
\]

when the normalized cutoff time \( x_m \geq \sqrt{2L(\varepsilon_x^{(1)})} \). Here, \( L(\varepsilon) \) is the tail function of the cooling function \( g(h) \) defined in Definition 2.

Proof. We have,

\[
|D_r^{(x_m)}(E) - D_r| \leq (\int_{x_m}^{\infty} + \int_{-\infty}^{-x_m}) dy |\tilde{p}(y)e^{-i\gamma y E}| \langle \psi_0 | e^{i\gamma y H} | \psi_0 \rangle |
\leq (\int_{x_m}^{\infty} + \int_{-\infty}^{-x_m}) dy \tilde{p}(y) =: \varepsilon_x^{(1)}.
\]

Recall that

\[
\tilde{p}(y) = \frac{1}{2} \int_{-\infty}^{\infty} p\left(\frac{z + y}{2}\right)p\left(\frac{z - y}{2}\right)dz
= \int_{-\infty}^{\infty} p(t - y)p(t)dt = [p \ast p](y).
\]

We can show the following relationship between \( \tilde{p}(y) \) and \( p(x) \) by the nonnegativity of \( p(x) \),

\[
\int_{-x_m}^{x_m} \tilde{p}(y)dy \geq \int_{-x_m/\sqrt{2}}^{x_m/\sqrt{2}} dx \int_{-x_m/\sqrt{2}}^{x_m/\sqrt{2}} dx' p(x)p(x')
\geq (1 - L^{-1}(\frac{x_m}{\sqrt{2}}))^2
\geq 1 - 2L^{-1}(\frac{x_m}{\sqrt{2}}),
\]

where \( \varepsilon = L(x_m) = 1 - \int_{-x_m}^{x_m} p(x)dx \).

Therefore,

\[
\varepsilon_x^{(1)} = 1 - \int_{-x_m}^{x_m} \tilde{p}(y)dy \leq 2L^{-1}(\frac{x_m}{\sqrt{2}}),
\]

\[
\Rightarrow x_m \geq \sqrt{2L(\varepsilon_x^{(1)})/2}.
\]

\[\square\]

C. Finite sampling number

Now, we consider the statistical fluctuation when estimating the values of \( D_r^{(x_m)}(E) \) using Eq. (102). In the single-shot version of Hadamard test in Algorithm 2, we can describe the identical and identically distributed (i.i.d.) single-round estimators for \( D_r^{(x_m)}(E) \) as \( \{\hat{d}_p(E)\}_{p=1}^{N_M} \). Each single-round estimator \( \hat{d}_p(E) \) is a random variable defined as follows,

\[
\hat{d}_p = \begin{cases} 
\text{Re}(2e^{-i\gamma y E}), & (b, a) = (0, 0), y : |y| \leq x_m \\
\text{Re}(-2e^{-i\gamma y E}), & (b, a) = (0, 1), y : |y| \leq x_m \\
\text{Re}(2i e^{-i\gamma y E}), & (b, a) = (1, 0), y : |y| \leq x_m \\
\text{Re}(-2i e^{-i\gamma y E}), & (b, a) = (1, 1), y : |y| \leq x_m \\
0, & y : |y| > x_m.
\end{cases}
\]
The final estimation of $D^{(x_m)}_{\tau}(E)$ is given by

$$
\hat{D}^{(x_m)}_{\tau}(E) = \frac{1}{N_M} \sum_{p=1}^{N_M} \hat{d}_p(E). \tag{109}
$$

Based on Eq. (102), we know that

$$
\mathbb{E}_{(y,b,a)} \left( \hat{D}^{(x_m)}_{\tau} \right) = D^{(x_m)}_{\tau}. \tag{110}
$$

To analyze the statistical fluctuation, we apply the Hoeffding bound.

**Lemma 1** (Hoeffding bound). For $n$ independent random variables $\{\hat{X}_i\}_{i=1}^n$ which are bounded by $[a,b]$, the average value

$$
\hat{X} := \frac{1}{n} \sum_{i=1}^{n} \hat{X}_i, \tag{111}
$$

satisfies

$$
\Pr( |\hat{X} - \mathbb{E}(\hat{X})| \geq \varepsilon ) \leq 2 \exp \left( -\frac{2n\varepsilon^2}{(b-a)^2} \right). \tag{112}
$$

**Proposition 5** (Error of the normalization factor introduced by finite sampling number). The estimator $\hat{D}^{(x_m)}_{\tau}(E)$ defined in Eq. (109) is related to the truncated estimation $\tilde{D}^{(x_m)}_{\tau}(E)$ defined in Eq. (102) by

$$
|\tilde{D}^{(x_m)}_{\tau}(E) - D^{(x_m)}_{\tau}(E)| \leq \varepsilon_n^{(1)}, \forall E \in \mathbb{R}, \tag{113}
$$

with a failure probability $\delta^{(1)} := 2 \exp \left( -2K(\varepsilon_n^{(1)})^2/16 \right)$, when the sample number $N_M \geq K(\varepsilon_n^{(1)})^{-2}$.

**Proof.** Note that, the estimators $\{\hat{d}_p(E)\}_{p=1}^{N_M}$ are independent, whose values are bounded by $[-2,2]$. Using the Hoeffding bound in Lemma 1 with the bound $[-2,2]$, we finish the proof. \qed

### D. Accuracy of the eigenenergy estimation

Ideally, the eigenenergy $E_j$ satisfies,

$$
E_j = \arg \max_{E \in [E^L_j, E^U_j]} D_{\tau}(E). \tag{114}
$$

In practice, what we can solve is the following problem,

$$
\hat{E}_j = \arg \max_{E \in [E^L_j, E^U_j]} \hat{D}^{(x_m)}_{\tau}(E). \tag{115}
$$

We now show that when $[E^L_j, E^U_j] \subset [E_j - \frac{\kappa}{2}, E_j + \frac{\kappa}{2}]$, $\hat{E}_j$ is a good estimation of $E_j$ with precision $\kappa$ when the finite imaginary time $\tau = \mathcal{O}(\kappa^{-1})$, finite normalized cutoff time $x_m = \mathcal{O}(L(p_j))$, and sample number $N_M = \mathcal{O}(p_j^{-2})$.

**Theorem 3** (Accuracy of the eigenenergy estimation). When $[E^L_j, E^U_j] \subset [E_j - \frac{\kappa}{2}, E_j + \frac{\kappa}{2}]$, the eigenenergy estimation $\hat{E}_j$ defined in Eq. (115) is related to the eigenenergy $E_j$ by

$$
|\hat{E}_j - E_j| \leq \kappa, \tag{116}
$$

with a failure probability of $2\delta^{(1)} = 4 \exp(-K/8)$, when the finite imaginary time $\tau \geq \frac{1}{8}g^{-1}(\frac{1-g(1)}{6}p_j)$, the normalized cutoff time $x_m \geq \sqrt{2L \left( \frac{1-g(1)}{6}p_j \right)}$, and the sample number $N_M \geq \frac{9K}{(1-g(1))\tau p_j^{-2}}$. 

Proof. From Eq. (115) we know that,
\[ \hat{D}_{\tau}^{(x_m)}(\hat{E}_j) \geq \hat{D}_{\tau}^{(x_m)}(E_j). \] (117)
Now, we take \( x_m = \sqrt{2L(\tau)} \) and \( N_M = (\varepsilon_n p_j)^{-2}K \). Using Proposition 4 and 5, we have
\[ |\hat{D}_{\tau}^{(x_m)}(\hat{E}_j) - D_{\tau}(\hat{E}_j)| \leq |D_{\tau}^{(x_m)}(\hat{E}_j) - D_{\tau}(\hat{E}_j)| + |\hat{D}_{\tau}^{(x_m)}(\hat{E}_j) - D_{\tau}(\hat{E}_j)| \leq p_j(\varepsilon_x + \varepsilon_n), \]
\[ |\hat{D}_{\tau}^{(x_m)}(E_j) - D_{\tau}(E_j)| \leq |D_{\tau}^{(x_m)}(E_j) - D_{\tau}(E_j)| + |\hat{D}_{\tau}^{(x_m)}(E_j) - D_{\tau}(E_j)| \leq p_j(\varepsilon_x + \varepsilon_n) \] (118)
each with a failure probability \( \delta^{(1)} \). Furthermore, when \( \tau = \frac{2}{\Delta} g^{-1}(\frac{p_j}{2}) \), from Proposition 3 we have
\[ |D_{\tau}(E_j) - p_j g(\tau(E_j - \hat{E}_j))^2| \leq p_j \varepsilon_x, \]
\[ |D_{\tau}(E_j) - p_j| \leq p_j \varepsilon_x. \] (119)
Combine Eqs. (117), (118), and (119), we have
\[ p_j \leq D_{\tau}(E_j) + \varepsilon_x \leq \hat{D}_{\tau}^{(x_m)}(E_j) + p_j(\varepsilon_x + \varepsilon_n) + p_j \varepsilon_x \]
\[ \leq \hat{D}_{\tau}^{(x_m)}(\hat{E}_j) + p_j(\varepsilon_x + \varepsilon_n) + p_j \varepsilon_x \]
\[ \leq D_{\tau}(\hat{E}_j) + 2p_j(\varepsilon_x + \varepsilon_n) + p_j \varepsilon_x, \]
\[ \leq p_j g(\tau(E_j - \hat{E}_j))^2 + 2p_j(\varepsilon_x + \varepsilon_n), \] (120)
with a failure probability \( 2\delta^{(1)} \). Therefore,
\[ g(\tau(E_j - \hat{E}_j))^2 \geq 1 - 2(\varepsilon_x + \varepsilon_n). \] (121)
Denote \( \kappa := \hat{E}_j - E_j \), we further simplify the expression above,
\[ \kappa \leq \frac{1}{\tau} g^{-1}(\sqrt{1 - 2(\varepsilon_x + \varepsilon_n)}) \]
\[ \leq \frac{1}{\tau} g^{-1}(1 - (\varepsilon_x + \varepsilon_n)) \]
\[ = \frac{1}{\tau} g^{-1} \left( 1 - \frac{2}{p_j} g(\frac{\Delta}{2}) - \frac{2}{p_j} L^{-1}(\frac{x_m}{\sqrt{2}}) - \sqrt{\frac{K}{p_j^2 N_M}} \right) \] (122)
In the second inequality, we use the property that \( g^{-1}(p) \) is a decreasing function when \( p > 0 \). Therefore, the following requirement is sufficient to make sure the eigenenergy error is smaller than \( \kappa \),
\[ \tau \geq \frac{1}{\kappa}, \]
\[ \frac{2}{p_j} g(\frac{\Delta}{2}) + \frac{2}{p_j} L^{-1}(\frac{x_m}{\sqrt{2}}) + \sqrt{\frac{K}{p_j^2 N_M}} \leq 1 - g(1). \] (123)
Let \( \frac{2}{p_j} g(\frac{\Delta}{2}) = \frac{2}{p_j} L^{-1}(\frac{x_m}{\sqrt{2}}) = \sqrt{\frac{K}{p_j^2 N_M}} \), we have
\[ \tau \geq 2 \Delta g^{-1}(\frac{1 - g(1)}{6} p_j), \]
\[ x_m \geq \sqrt{2L} \left( \frac{1 - g(1)}{6} p_j \right), \]
\[ N_M \geq K \left( \frac{3}{1 - g(1)} \right)^2 \frac{1}{p_j^2}. \] (124)
In the above derivation, we require \( \tau \geq \frac{1}{\kappa} \) and \( \tau \geq \frac{2}{\Delta} g^{-1}(\frac{1 - g(1)}{6} p_j) \). In practice, \( \kappa \ll \frac{\Delta}{2} \). Then it suffices to have
\[ \tau \geq \frac{1}{\kappa} g^{-1}(\frac{1 - g(1)}{6} p_j). \]
\[ \] (125)
From Theorem 3 we can see that, the (maximal) circuit depth and sample complexity of Algorithm 2 are \( \tau x_m = O(\kappa^{-1} g^{-1}(p_j) L(p_j)) \) and \( N_M = O(p_j^{-2}) \), respectively, for a given accuracy \( \kappa \) and initial state overlap \( p_j \). This achieves the Heisenberg limit \( \frac{1}{\kappa} \) for the eigenenergy searching.
V. ERROR AND RESOURCE REQUIREMENT ANALYSIS FOR THE OBSERVABLE ESTIMATION

In this section, we study the estimation error of the observable value \( \langle O \rangle \) in Algorithm 5 under the finite circuit depth and sampling number. Based on the error dependence, we estimate the resource requirements (i.e. circuit depth and sample number) of the cooling algorithm.

Similar to Sec. IV, we limit our analysis to the positive and even cooling function such that \( g(h) = g(-h) \). For simplicity, we first assume that we have already obtained a precise eigenenergy estimation \( E_j \) of the \( j \)-th eigenstate following Algorithm 2.

In this section, we study the estimation error of the observable value \( \langle O \rangle \) in Algorithm 5 under the finite circuit depth and sampling number. Based on the error dependence, we estimate the resource requirements (i.e. circuit depth and sample number) of the cooling algorithm.

In Fig. 5 we summarize our error analysis. In Sec. VA, we consider the effect of finite imaginary time \( \tau \); in Sec. VB, we bound the estimation error caused by the normalized cutoff time \( x_m \). The two factors above determine the actual maximum evolution time. In Sec. VC, based on the measurement using Hadamard test, we determine the statistical error caused by finite sampling number \( N_M \). Finally, we summarize the finite sampling effect in Sec.VD and show the dependence of circuit depth \( t_m \) and sampling number \( N_M \) with respect to the initial state \( |\psi_0\rangle \), Hamiltonian \( H \) and the observable \( O \).

Later on, in Sec. VE, we will discuss the effect of the eigenenergy estimation error \( \kappa \) on the observable estimation. When \( \kappa \ll \min\{|E_j - E_{j-1}|, |E_j - E_{j+1}|\} \), the error \( \kappa \) will introduce negligible effect on the accuracy of the observable estimation. However, it will introduce a \( \tau \)-dependent factor on the normalization factor, which causes a higher sampling cost.

\[
\langle O \rangle = \frac{N(O)}{D},
\]

Sec. IVA and VA finite \( \tau \)

\[
\langle O \rangle = \frac{N_\tau(O)}{D_\tau},
\]

Sec. IVB and VB finite \( x_m \)

\[
\langle O \rangle_{\tau}^{(x_m)} = \frac{N_{\tau}^{(x_m)}(O)}{D_{\tau}^{(x_m)}},
\]

Sec. IV C and VC finite \( N_M \)

\[
\langle \tilde{O} \rangle_{\tau}^{(x_m)} = \frac{\tilde{N}_{\tau}^{(x_m)}(O)}{\tilde{D}_{\tau}^{(x_m)}},
\]

When \( \tau = \max\{\frac{1}{\Delta} g^{-1}(\epsilon_\tau^{(1)}/2), \frac{1}{\Delta} g^{-1}(\epsilon_\tau^{(2)}/2)\} \):

\[
|D_\tau - D| \leq \epsilon_\tau^{(1)},
\]

\[
|N_\tau(O) - N(O)| \leq \epsilon_\tau^{(2)} \|O\|_\infty,
\]

When \( x_m = \max\{\sqrt{2} L(\epsilon_x^{(1)/2}), L(\epsilon_x^{(2)/2})\} \):

\[
|D_{\tau}^{(x_m)} - D| \leq \epsilon_x^{(1)}
\]

\[
|N_{\tau}^{(x_m)}(O) - N_\tau(O)| \leq \epsilon_x^{(2)} \|O\|_\infty.
\]

When \( N_M = \max\left\{\left(\frac{\epsilon_\eta^{(1)}}{\epsilon_\eta^{(2)}}\right)^2, \left(\frac{\epsilon_\eta^{(2)}}{\epsilon_\eta^{(1)}}\right)^2\right\} \):

\[
|\tilde{D}_{\tau}^{(x_m)} - D_{\tau}^{(x_m)}| \leq \epsilon_\eta^{(1)}
\]

\[
|\tilde{N}_{\tau}^{(x_m)}(O) - \tilde{N}_{\tau}^{(x_m)}(O)| \leq \epsilon_\eta^{(2)} \|O\|_1.
\]

FIG. 5. Summary of the error analysis of the observable estimation. Started from the ideal observation value \( \langle O \rangle \), we sequentially study the effect of finite imaginary time \( \tau \), finite normalized cutoff time \( x_m \) and finite sampling number \( N_M \).

A. Finite imaginary evolution time

Recall that

\[
\langle O \rangle = \langle u_j | O | u_j \rangle = \frac{N(O)}{D},
\]

where

\[
D = D(E_j) = \langle \psi_0 | \hat{P}_j | \psi_0 \rangle = p_j;
\]

\[
N(O) = \langle \psi | \hat{P}_j O \hat{P}_j | \psi \rangle = p_j \langle O \rangle,
\]

are, respectively, the normalization factor and the unnormalized expectation value. Here, \( \hat{P}_j := |u_j\rangle \langle u_j| \) is the projector to the \( j \)-th eigenstate.

When the imaginary time \( \tau \) is finite, we have the approximated value

\[
\langle O \rangle_{\tau} = \frac{N_{\tau}(O)}{D_{\tau}},
\]

where

\[
D_{\tau} = D(E_j) = \langle \psi_0 | \hat{P}_{j_{\tau}} | \psi_0 \rangle = p_{j_{\tau}};
\]

\[
N_{\tau}(O) = \langle \psi | \hat{P}_{j_{\tau}} O \hat{P}_{j_{\tau}} | \psi \rangle = p_{j_{\tau}} \langle O \rangle_{\tau},
\]
where
\[ D_{\tau} = \langle \psi_0 | g^2(\tau (H - E_j)) | \psi_0 \rangle, \]
\[ N_{\tau}(O) = \langle \psi_0 | g(\tau (H - E_j)) O g(\tau (H - E_j)) | \psi_0 \rangle. \]  
(128)

are, respectively, the normalization factor and the unnormalized expectation value with finite \( \tau \).

In Proposition 3 in Sec. IV A, we have analyzed the distance between \( D_{\tau}(E) \) and \( D(E) \) when \( |E - E_j| \leq \frac{\Delta}{2} \). Similarly, we now bound the distance between \( N_{\tau}(O) \) and \( N(O) \).

**Proposition 6** (Error of the unnormalized expectation value introduced by finite imaginary time). The unnormalized expectation value \( N_{\tau}(O) \) defined in Eq. (128) is related to the ideal \( N(O) \) defined in Eq. (126) by
\[ |N_{\tau}(O) - N(O)| \leq \|O\|_{\infty} \varepsilon_{\tau}(2), \]  
(129)
when the imaginary time \( \tau \geq \frac{1}{\Delta} g^{-1}(\frac{\varepsilon_{\tau}(2)}{2}) \). Here, \( \|O\|_{\infty} \) is the spectral norm of \( O \).

**Proof.** To bound the distance between \( N_{\tau}(O) \) and \( N(O) \), we first bound the distance between \( \hat{P}_j \) and \( g(\tau (H - E_j)) \).

We have,
\[ g(\tau (H - E_j)) - \hat{P}_j = \sum_{i \neq j} g(\tau (E_i - E_j)) |u_i \rangle \langle u_i| . \]  
(130)

Denote \( \Delta \) to be the known lower bound of \( \min \Delta_{j,j+1}, \Delta_{j,j-1} \). If we choose \( \tau \) to be \( \tau \geq \frac{1}{\Delta} g^{-1}(\frac{\varepsilon_{\tau}(2)}{2}) \), we have
\[ g(\tau (E_i - E_j)) = g \left( \frac{|E_i - E_j|}{\Delta} g^{-1}(\frac{1}{2} \varepsilon_{\tau}(2)) \right) \]
\[ \leq g \left( g^{-1}(\frac{1}{2} \varepsilon_{\tau}(2)) \right) \]
\[ = \frac{1}{2} \varepsilon_{\tau}(2). \]  
(131)

In the second line, we use the definition of \( g(h) \) that \( g(h') \leq g(h) \) for all \( h' > h > 0 \). Based on Eq. (130) and Eq. (131), we have
\[ \|g(\tau (H - E_j)) - \hat{P}_j\|_{\infty} \leq \frac{1}{2} \varepsilon_{\tau}(2), \]  
(132)
when \( \tau \geq \frac{1}{\Delta} g^{-1}(\frac{\varepsilon_{\tau}(2)}{2}) \).

Now, we use Eq. (132) to bound the difference between \( N_{\tau}(O) \) and \( N(O) \). When \( \tau \geq \frac{1}{\Delta} g^{-1}(\frac{\varepsilon_{\tau}(2)}{2}) \), we then have,
\[ |N_{\tau}(O) - N(O)| = |\langle \psi_0 | g(\tau (H - E_j)) O g(\tau (H - E_j)) | \psi_0 \rangle - \langle \psi_0 | \hat{P}_j O \hat{P}_j | \psi_0 \rangle| \]
\[ \leq \left| \langle \psi_0 | g(\tau (H - E_j)) O g(\tau (H - E_j)) | \psi_0 \rangle - \langle \psi_0 | \hat{P}_j O \hat{P}_j | \psi_0 \rangle \right| \]
\[ + \left| \langle \psi_0 | \hat{P}_j O g(\tau (H - E_j)) | \psi_0 \rangle - \langle \psi_0 | \hat{P}_j \hat{O} | \psi_0 \rangle \right| \]
\[ \leq 2 \|g(\tau (H - E_j)) - \hat{P}_j\|_{\infty} \|O\|_{\infty} \varepsilon_{\tau}(2) \]  
(133)

B. Finite normalized cutoff time

In Proposition 4, we bound the estimation error of \( D_{\tau}(O) \) caused by finite \( x_m \). Follow similar methods, we bound the estimation errors of \( N_{\tau}(O) \) caused by finite \( x_m \). The estimation formula is given in Eq. (75),
\[ N^{(x_m)}_{\tau}(O) = \|O\|_1 \int_{-x_m}^{x_m} dx \int_{-x_m}^{x_m} dx' P(x)p(x') \sum_{l \in \mathcal{P}} \text{Pr}(l)e^{-i\tau(x-x')} E_l \langle \psi_0 | e^{-i\tau x'H} P_l e^{i\tau x'H} | \psi_0 \rangle , \]  
(134)
\[ O = \sum_{l \in P_n} o_l P_l = \|O\|_1 \sum_{l \in P_n} \Pr_O(l) P_l, \]  
(135)

\[ \|O\|_1 \text{ is the } l_1\text{-norm of Pauli coefficients of } O, \]

\[ \|O\|_1 = \sum_{l \in P_n} o_l. \]  
(136)

Note that, the coefficients \( \{o_l\}_l \) are all positive, since we put the signs into the Pauli matrices \( \{P_l\} \). The probability distribution \( \Pr_O(l) \) is defined to be,

\[ \Pr_O(l) = \frac{o_l}{\sum_{l \in P_n} o_l} = \frac{1}{\|O\|_1 o_l}. \]  
(137)

**Proposition 7** (Error of the unnormalized expectation value introduced by finite normalized cutoff time). The truncated estimation \( N_{\tau}^{(x_m)}(O) \) defined in Eq. (134) is related to the normalization factor \( N_{\tau}(O) \) defined in Eq. (128) by

\[ |N_{\tau}^{(x_m)}(O) - N_{\tau}(O)| \leq \|O\|_\infty \varepsilon_x^{(2)}, \]  
(138)

when the normalized cutoff time \( x_m \geq L(\frac{\varepsilon_x^{(2)}}{2}) \). Here, \( L(\varepsilon) \) is the tail function of the cooling function \( g(h) \) defined in Definition 2.

**Proof.** We have,

\[ |N_{\tau}^{(x_m)}(O) - N_{\tau}(O)| \leq \iint_{\bar{S}} dx dx' p(x)p(x') |e^{-i\tau(x-x')E_1} \langle \psi_0|e^{-i\tau x'H}Oe^{i\tau x'H} |\psi_0\rangle| \]

\[ \leq \iint_{\bar{S}} dx dx' p(x)p(x') \left| \langle \psi_0|e^{-i\tau x'H}Oe^{i\tau x'H} |\psi_0\rangle \right| \]

\[ \leq \|O\|_\infty \iint_{\bar{S}} dx dx' p(x)p(x') =: \|O\|_\infty \varepsilon_x^{(2)}, \]  
(139)

where \( \bar{S} \) implies the complement area of \( S : \{(x, x')|x| \leq x_m, |x'| \leq x_m\} \).

The error term \( \varepsilon_x^{(2)} \) is related to the tail probability by

\[ \varepsilon_x^{(2)} = 1 - \int_{\bar{S}} dx dx' p(x)p(x') = 1 - (1 - L^{-1}(x_m))^2 \leq 2L^{-1}(x_m) \]

\[ \Rightarrow \quad x_m \geq L(\frac{\varepsilon_x^{(2)}}{2}). \]  
(140)

Therefore, when we choose the \( x_m \) to be \( L(\varepsilon_x^{(2)}) \), we can achieve the estimation of \( N_{\tau}(O) \) with a precession of \( \|O\|_\infty \varepsilon_x^{(2)} \).

\[ \square \]

**C. Finite sample number**

In Proposition 4, we bound the estimation error of \( D_{\tau}^{(x_m)}(O) \) caused by finite sampling error \( N_M \). Follow similar methods, we bound the statistical error of \( N_{\tau}^{(x_m)}(O) \).

In the single-shot version of Hadamard test in Algorithm 5, we can describe these identical and identically distributed (i.i.d.) single-round estimators for \( N_{\tau}^{(x_m)}(O) \) as \( \{\hat{n}_q\}_{q=1}^{N_M} \). Each single-round estimator \( \hat{n}_q \) is a random variable defined as follows,

\[ \hat{n}_q(O) = \begin{cases} 
\text{Re}(2\|O\|_1 e^{-i\tau(x-x')E_1}), & (b, a) = (0, 0), (x, x') : |x| \leq x_m, |x'| \leq x_m \\
\text{Re(-2\|O\|_1 e^{-i\tau(x-x')E_1}),} & (b, a) = (0, 1), (x, x') : |x| \leq x_m, |x'| \leq x_m \\
\text{Re(2\|O\|_1 e^{-i\tau(x-x')E_1}),} & (b, a) = (1, 0), (x, x') : |x| \leq x_m, |x'| \leq x_m \\
\text{Re(-2\|O\|_1 e^{-i\tau(x-x')E_1}),} & (b, a) = (1, 1), (x, x') : |x| \leq x_m, |x'| \leq x_m \\
0, & (x, x') : |x| > x_m \text{ or } |x'| > x_m.
\end{cases} \]  
(141)
The final estimation of $N_{\tau}(x_m)(O)$ is given by

$$\hat{N}_{\tau}(x_m)(O) = \frac{1}{N_M} \sum_{q=1}^{N_M} \hat{n}_q(O).$$

(142)

Based on Eq. (134), we know that

$$E_{(x,x',p,b,a)} (\hat{N}_{\tau}(x_m)(O)) = N_{\tau}(x_m)(O).$$

(143)

Now we bound the statistical error of $\hat{N}_{\tau}(x_m)(O)$ using Lemma 1.

**Proposition 8** (Error of the observable expectation value introduced by finite sampling number). The estimator $\hat{N}_{\tau}(x_m)(O)$ defined in Eq. (142) is related to the truncated estimation $\hat{N}_{\tau}(x_m)(O)$ defined in Eq. (134) by

$$|\hat{N}_{\tau}(x_m)(O) - N_{\tau}(x_m)(O)| \leq \|O\|_1 \varepsilon_{n}^{(2)},$$

(144)

with a failure probability $\delta^{(2)} := 2 \exp \left(-2K(\varepsilon_{n}^{(2)})^2/16\right)$, when the sample number $N_M \geq K(\varepsilon_{n}^{(2)})^{-2}$. Here, $\|O\|_1 = \sum_i |a_i|$ is the sum of the Pauli coefficients of $O$.

Proof. Note that, the estimators $\{\hat{n}_q(O)\}_{q=1}^{N_M}$ are independent, whose values are bounded by $[-2\|O\|_1, 2\|O\|_1]$. Here, $\|O\|_1 = \sum_i |a_i|$. Using the Hoeffding bound in Lemma 1 with the bound $[-2\|O\|_1, 2\|O\|_1]$, we finish the proof. \(\square\)

**D. Accuracy of the observable estimation**

The final observable estimation is given by

$$\langle \hat{O} \rangle_{\tau}(x_m) = \frac{\hat{N}_{\tau}(x_m)}{O_{\tau}(x_m)}.$$

(145)

In Propositions 3, 4, 5, 6, 7, and 8, we estimate the errors caused by finite $\tau$, $x_m$, and $N_M$ on $D_{\tau}$ and $N_{\tau}(O)$, respectively. Now we combine the results together.

**Proposition 9.** For the initial state $|\psi_0\rangle$, target eigenstate $|u_j\rangle$ and cooling function $g(h)$, if we set the imaginary time $\tau$, normalized cutoff time $x_m$, and sample number $N$ to be

1. $\tau = \frac{1}{N} \max \{g^{-1}(\frac{\varepsilon^{(1)}}{2}), g^{-1}(\frac{\varepsilon^{(2)}}{2})\};$
2. $x_m = \max \{\sqrt{2L(\frac{\varepsilon^{(1)}}{2}), L(\frac{\varepsilon^{(2)}}{2})}\};$
3. $N_M = \max \{K(\varepsilon^{(1)})^{-2}, K(\varepsilon^{(2)})^{-2}\};$

then the estimated observable expectation value $\langle \hat{O} \rangle_{\tau}(x_m)$ defined in Eq. (145) in Algorithm 5 is related to the real observable expectation $\langle O \rangle$ by,

$$|\langle \hat{O} \rangle_{\tau}(x_m) - \langle O \rangle | \leq p_1^{-1}(\|O\|_1 + 1)(\varepsilon^{(1)}_\tau + \varepsilon^{(1)}_x + \varepsilon^{(1)}_n) + p_2^{-1}(\varepsilon^{(2)}_\tau)\|O\|_\infty + \varepsilon^{(2)}_x\|O\|_\infty + \varepsilon^{(2)}_n\|O\|_1),$$

(146)

with a failure probability $\delta^{(1)} + \delta^{(2)} = 4e^{-\frac{K}{2}}$.

Proof. Following Propositions 3, 4, 5, 6, 7, and 8, when we choose $\tau, x_m, N_M$ to be the values mentioned above, then the following bound holds at the same time

$$|\hat{D}(x_m) - D| \leq \varepsilon^{(1)}_\tau + \varepsilon^{(1)}_x + \varepsilon^{(1)}_n,$$

$$|\hat{N}(x_m)(O) - N(O)| \leq (\varepsilon^{(2)}_\tau)\|O\|_\infty + \varepsilon^{(2)}_x\|O\|_\infty + \varepsilon^{(2)}_n\|O\|_1),$$

(147)

with a failure probability $\delta(K) = 4e^{-\frac{K}{2}}$. To simplify the notation, we denote $\varepsilon^{(1)} := \varepsilon^{(1)}_\tau + \varepsilon^{(1)}_x + \varepsilon^{(1)}_n$ and $\varepsilon^{(2)}(O) := \varepsilon^{(2)}_\tau\|O\|_\infty + \varepsilon^{(2)}_x\|O\|_\infty + \varepsilon^{(2)}_n\|O\|_1$. 
Based on Eq. (147), we bound the difference between the quotients \( \langle \hat{O} \rangle_{\tau}^{(x_m)} \) and \( \langle O \rangle \),

\[
| \langle O \rangle - \langle \hat{O} \rangle_{\tau}^{(x_m)} | = \left| \frac{N(O)}{D} - \frac{\hat{N}_{\tau}^{(x_m)}(O)}{\hat{D}_{\tau}^{(x_m)}} \right| \\
= \left| \frac{N(O)\hat{D}_{\tau}^{(x_m)} - D\hat{N}_{\tau}^{(x_m)}(O)}{\hat{D}_{\tau}^{(x_m)}} \right| \\
\leq \frac{|N(O)(D + \varepsilon^{(1)}) - D(N(O) - \varepsilon^{(2)}(O))|}{D(D - \varepsilon^{(1)})} \\
= \frac{|N(O)|\varepsilon^{(1)} + D\varepsilon^{(2)}(O)|}{D^2 - D\varepsilon^{(1)}} \\
\leq \frac{(N(O) + D)\varepsilon^{(1)} + D\varepsilon^{(2)}(O)}{D^2} \\
= p_j^{-1}(\langle O \rangle + 1)(\varepsilon^{(1)}_r + \varepsilon^{(2)}_x + \varepsilon^{(1)}_n) + p_j^{-1}(\varepsilon^{(2)}_r ||O||_\infty + \varepsilon^{(2)}_x ||O||_\infty + \varepsilon^{(2)}_n ||O||_1).
\]

In the fifth line, we use the fact that \( N(O)\varepsilon^{(1)} \) and \( D\varepsilon^{(2)}(O) \) are much smaller than \( D^2 = p_j^2 \).

Based on Proposition 9, we can estimate the circuit depth and sample complexity for the observable estimation task.

**Theorem 4** (Accuracy of the observable estimation). The expectation value estimation \( \langle \hat{O} \rangle_{\tau}^{(x_m)} \) defined in Eq. (145) is related to the real observable expectation value \( \langle O \rangle \) by

\[
| \langle O \rangle - \langle \hat{O} \rangle_{\tau}^{(x_m)} | \leq \varepsilon \left( \frac{1}{2}(\langle O \rangle + 1) + \frac{1}{3}||O||_\infty + \frac{1}{6}||O||_1 \right)
\]

(149)

with a failure probability of \( \delta^{(1)} + \delta^{(2)} = 4 \exp(-K/8) \), when the finite imaginary time \( \tau \geq \frac{1}{2}g^{-1}(\frac{2p_j}{12}) \), the normalized cutoff time \( x_m \geq \sqrt{2L(\frac{2p_j}{12})} \), and the sample number \( N_M \geq K(\frac{2p_j}{12})^{-2} \). Here, \( ||O||_1 = \sum_i |o_i| \) is the sum of the Pauli coefficients of \( O \), \( \Delta \) is a known lower bound of \( \min\{E_j - E_{j-1}, |E_j - E_{j+1}|\} \).

**Proof.** From Proposition 9 we know that, if we set \( \tau = \frac{1}{2} \max\{g^{-1}(\frac{2p_j}{12}), g^{-1}(\frac{2p_j}{12})\} \) \( x_m = \max\{\sqrt{2L(\frac{2p_j}{12})}, L(\frac{2p_j}{12})\} \) and \( N_M = \max\{K(\frac{2p_j}{12})^{-2}, K(\frac{2p_j}{12})^{-2}\} \), then

\[
| \langle O \rangle - \langle \hat{O} \rangle_{\tau}^{(x_m)} | \leq p_j^{-1}(\langle O \rangle + 1)(\varepsilon^{(1)}_r + \varepsilon^{(2)}_x + \varepsilon^{(1)}_n) + p_j^{-1}(\varepsilon^{(2)}_r ||O||_\infty + \varepsilon^{(2)}_x ||O||_\infty + \varepsilon^{(2)}_n ||O||_1) \\
\leq p_j^{-1}(||O||_1 + 1)(\varepsilon^{(1)}_r + \varepsilon^{(2)}_x + \varepsilon^{(1)}_n) + p_j^{-1}||O||_1(\varepsilon^{(2)}_r + \varepsilon^{(2)}_x + \varepsilon^{(2)}_n)
\]

(150)

If we set

\[
\varepsilon^{(1)}_r = \varepsilon^{(2)}_r = \varepsilon^{(1)}_x = \varepsilon^{(2)}_x = \varepsilon^{(1)}_n = \varepsilon^{(2)}_n = \frac{\varepsilon p_j}{6},
\]

(151)

which corresponds to

\[
\tau \geq \frac{1}{2}g^{-1}(\frac{\varepsilon p_j}{12}) \\
x_m \geq \sqrt{2L(\frac{\varepsilon p_j}{12})},
\]

(152)

then Eq. (150) becomes,

\[
| \langle \hat{O} \rangle_{\tau}^{(x_m)} - \langle O \rangle | \leq \varepsilon \left( \frac{1}{2}(\langle O \rangle + 1) + \frac{1}{3}||O||_\infty + \frac{1}{6}||O||_1 \right) \\
\leq \varepsilon(||O||_1 + 1).
\]

(153)
From Theorem 4 we can see that, the circuit depth and sample complexity of the observable estimation are \( \tau x_m = \mathcal{O}(\frac{1}{2}g^{-1}(\varepsilon_2) L(\varepsilon_1)) \) and \( N_M = \mathcal{O}(\varepsilon^{-2} I_2^{-1}) \), respectively.

As is shown in Eqs. (23), (43), (52), and (61), the common cooling functions satisfy
\[
g^{-1}(\varepsilon) = \mathcal{O}(\log(\varepsilon^{-1})),
\]
that is, with a logarithm increasing \( \tau \), we can suppress the error term caused by finite imaginary time evolution. On the other hand, for the common cooling functions,
\[
L(\varepsilon) = \mathcal{O}(\text{poly}(\varepsilon^{-1})).
\]
This implies that the circuit depth of a general cooling function is \( \mathcal{O}(\text{poly}(\varepsilon^{-1})) \log(\varepsilon^{-1}) \). For the Gaussian cooling or secant hyperbolic function we can further reduce it to \( \mathcal{O}(\text{poly}(\log(\varepsilon^{-1})) \), which is exponentially better than the phase estimation algorithm.

E. Effect of eigenenergy estimation error on the observable estimation

Now, we analyze the effect of the energy estimation error \( \kappa \). As a reasonable assumption, we assume that \( |\kappa| < \frac{\Delta}{2} \), i.e., the estimated eigenenergy \( \hat{E}_j \) satisfies \( |\hat{E}_j - E_j| \leq \frac{\Delta}{2} \).

The normalization factor \( D_x \) and unnormalized observation value \( N_x(O) \) will become
\[
D_x(\hat{E}_j) = \langle \psi_0| g^2(\tau(H - \hat{E}_j))|\psi_0 \rangle,
N_x(O; \hat{E}_j) = \langle \psi_0| g(\tau(H - \hat{E}_j))O g(\tau(H - \hat{E}_j))|\psi_0 \rangle.
\]

In practice, we will generate estimated values \( \hat{D}_x^{(\tau)}(\hat{E}_j) \) and \( \hat{N}_x^{(\tau)}(O; \hat{E}_j) \) following similar methods introduced in Sec. V B and Sec. V C. The final estimation of \( \langle O \rangle \) will be
\[
\langle O \rangle_{\tau}^{(\tau)}(\hat{E}_j) = \frac{\hat{N}_x^{(\tau)}(O; \hat{E}_j)}{\hat{D}_x^{(\tau)}(\hat{E}_j)}.
\]

To bound the distance between \( \langle O \rangle_{\tau}^{(\tau)}(\hat{E}_j) \) and \( \langle O \rangle \), we only need to relate \( \hat{N}_x^{(\tau)}(O; \hat{E}_j) \) and \( \hat{D}_x^{(\tau)}(\hat{E}_j) \) to \( N(O) \) and \( D \), respectively. The errors introduced by finite \( x_m \) and \( \hat{N}_M \) can be still analyzed based on Proposition 4, 5, 7, and 8. Furthermore, the finite \( \tau \) error in the normalization factor \( D(E) \) can be handled by Proposition 3. Therefore, we only need to generalize Proposition 6, following similar analysis in Proposition 3.

**Proposition 10** (Error of the unnormalized expectation value introduced by finite imaginary time with energy estimation inaccuracy). When \( E \in [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}] \), the finite \( \tau \) unnormalized expectation value \( N_x(O) \) defined in Eq. (156) is related to the ideal \( N(O) \) defined in Eq. (126) by
\[
|N_x(O; E) - g(\tau(E - E_j))^2 N(O)| \leq \|O\|_\infty \varepsilon_{\tau}^{(2)},
\]
when the imaginary time \( \tau \geq \frac{1}{2}g^{-1}(\varepsilon_{\tau}^{(2)}) \). Here, \( \|O\|_\infty \) is the spectral norm of \( O \).

**Proof.** From Eq. (100) in the proof in Proposition 3 we have
\[
\|g(\tau(H - E)) - g(\tau(E_j - E)) \hat{P}_j\| \leq \frac{\varepsilon_{\tau}^{(2)}}{2},
\]
when \( \tau \geq \frac{2}{\Delta}g^{-1}(\varepsilon_{\tau}^{(2)}) \).

Now, we use Eq. (159) to bound the difference between \( N_x(O; E) \) and \( g(\tau(E - E_j))^2 N(O) \),
\[
|N_x(O; E) - g(\tau(E - E_j))^2 N(O)| = | \langle \psi_0| g(\tau(H - E_j))O g(\tau(H - E_j))|\psi_0 \rangle - g(\tau(E - E_j))^2 \langle \psi_0| \hat{P}_j O \hat{P}_j |\psi_0 \rangle |
\leq | \langle \psi_0| g(\tau(H - E_j))O g(\tau(H - E_j))|\psi_0 \rangle - g(\tau(E - E_j)) \langle \psi_0| \hat{P}_j O g(\tau(H - E_j))|\psi_0 \rangle |
+ \langle \psi_0| \hat{P}_j O g(\tau(H - E_j))|\psi_0 \rangle - g(\tau(E - E_j)) \langle \psi_0| \hat{P}_j O \hat{P}_j |\psi_0 \rangle |
\leq \|g(\tau(H - E_j)) - \hat{P}_j\|_\infty \|O\|_\infty + g(\tau(E - E_j)) \|g(\tau(H - E_j)) - \hat{P}_j\|_\infty \|O\|_\infty
\leq 2 \|g(\tau(H - E_j)) - \hat{P}_j\|_\infty \|O\|_\infty
\leq \varepsilon_{\tau}^{(2)} \|O\|_\infty.
\]

(160)
Based on Propositions 3, 4, 5, 7, 8, and 10, we then have the following proposition, which is a modified version of Proposition 9.

**Proposition 11.** For the initial state \( |\psi_0\rangle \), target eigenstate \( |u_j\rangle \) and cooling function \( g(h) \), when the finite energy estimation error satisfies \( |\varepsilon| \leq \frac{\Delta}{2} \), if we set the imaginary time \( \tau \), normalized cutoff time \( x_m \), and sample number \( N \) to be

1. \( \tau = \frac{2}{\Delta} \max\{g^{-1}(\varepsilon_1^2, g^{-1}(\varepsilon_2^2))\} \);
2. \( x_m = \max\{\sqrt{2}L(\varepsilon_1^2, L(\varepsilon_2^2))\} \);
3. \( N_M = \max\{K(\varepsilon_1^2)^{-2}, K(\varepsilon_2^2)^{-2}\} \);

then the estimated observable expectation value \( \langle \hat{O}\rangle_{\tau}(x_m)(\hat{E}_j) \) defined in Eq. (145) in Algorithm 5 is related to the real observable expectation value \( \langle O \rangle \) by,

\[
|\langle \hat{O}\rangle_{\tau}(x_m) - \langle O \rangle| \leq \frac{1}{2}p_{p_j}^{-1}g(\tau\kappa)^{-2}(\langle O \rangle + 1)(\varepsilon_1^1 + \varepsilon_2^1 + \varepsilon_3^1) + p_{p_j}^{-1}g(\tau\kappa)^{-2}(\varepsilon_1^2||O||_{\infty} + \varepsilon_2^2||O||_{\infty} + \varepsilon_3^2||O||_1),
\]

with a failure probability \( \delta^{(1)} + \delta^{(2)} = 4e^{-\frac{K}{8}} \).

The proof of Proposition 11 is omitted since it is almost the same as the one in Proposition 9. We only need to replace \( N(O) \) and \( D \) to \( g(\tau\kappa)^2N(O) \) and \( g(\tau\kappa)^2D \), respectively.

Proposition 11 indicates that, when the eigenenergy is estimated to a precision of \( \kappa \), the observation error will become \( g(\tau\kappa)^2 \) times larger than the case when \( E_j \) is accurate. This will put an upper bound on the value of \( \tau \) when \( \tau \rightarrow \infty \), \( g(\tau\kappa)^2 \rightarrow 0 \).

Similar to Theorem 4, we have the following proposition.

**Proposition 12 (Accuracy of the observable estimation with small eigenenergy estimation error).** When the finite estimation error satisfies \( |\varepsilon| \leq \frac{\Delta}{2} \), the expectation value estimation \( \langle \hat{O}\rangle_{\tau}(x_m)(\hat{E}_j) \) defined in Eq. (157) is related to the real observable expectation value \( \langle O \rangle \) by

\[
|\langle \hat{O}\rangle_{\tau}(x_m) - \langle O \rangle| \leq \varepsilon\left(\frac{1}{2}(\langle O \rangle + 1) + \frac{1}{3}||O||_{\infty} + \frac{1}{6}||O||_1\right),
\]

with a failure probability \( \delta^{(1)} + \delta^{(2)} = 4e^{-K/8} \), when the finite imaginary time \( \tau \) satisfies

1. \( \tau \leq \tau_m \), where \( \tau_m = \mathcal{O}(1/\kappa) \) so that \( g(\tau_m\kappa) = \Omega(1) \);
2. \( \tau \geq \frac{2}{\Delta}g^{-1}\left(\frac{\varepsilon_{p_j}}{12}\right) \),

the normalized cutoff time \( x_m \geq \sqrt{2}L\left(\frac{\varepsilon_{p_j}g(\tau\kappa)^2}{12}\right) \), and the sample number \( N_M \geq K\left(\frac{\varepsilon_{p_j}g(\tau\kappa)^2}{6}\right)^{-2} \). Here, \( ||O||_1 = \sum_i |\alpha_i| \) is the sum of the Pauli coefficients of \( O \).

The proof is essentially the same as the one in Theorem 4. We only need to modify Eq. (151) to

\[
\varepsilon_1^1 = \varepsilon_2^1 = \varepsilon_3^1 = \varepsilon_1^2 = \varepsilon_2^2 = \varepsilon_3^2 = \frac{\varepsilon_{p_j}}{6}g(\tau\kappa)^2,
\]

so that Proposition 11 holds.

Now we discuss the requirement of energy estimation precision \( \kappa \) based on Proposition 12. To make sure that the observable \( O \) can be accurately estimated, the finite imaginary evolution time \( \tau \) must satisfy the following two requirements,

1. \( \tau \leq \tau_m \), where \( \tau_m = \mathcal{O}(1/\kappa) \) so that \( g(\tau_m\kappa) = \Omega(1) \);
2. \( \tau \geq \frac{2}{\Delta}g^{-1}\left(\frac{\varepsilon_{p_j}}{12}\right) \).

As a result, the value of \( \kappa \) should satisfy

\[
\kappa \leq \mathcal{O}\left(\frac{\Delta}{g^{-1}(\varepsilon_{p_j}/12)}\right).
\]

So that we are able to estimate the observable values with proper \( \tau \).
FIG. 6. We show the spectrum search of the 8-site anisotropic Heisenberg model using the Gaussian cooling function with different evolution time. The solid line shows the denominator of $D_\tau$ with different $\tau$ from $\tau = 0.9$ to 1.7. The cutoff is calculated according to the precision in the main text. We set $N_t = 10^5$ in the Monte Carlo sampling of the integral, and we use the expectation value for each sample that ignores measurement shot noise to keep consistent as that in the main text.

FIG. 7. Simulation error during the cooling process. The error is compared with the ideal expectation value given by $\langle \psi_0 | g(\tau(H - E_2))\dagger O g(\tau(H - E_2)) | \psi_0 \rangle$. The simulation error is below 0.05 under evolution.

VI. NUMERICAL SIMULATION

We show the spectrum search with different total evolution time $T = \tau x_m$ with the cutoff $x_m$ and imaginary evolution time $\tau$. Here, $\tau$ and $x_m$ are calculated according to the precision, as discussed in the above sections. The maximum imaginary evolution time $\tau$ ranges from $\tau = 0.9$ to 1.7. As shown in Figure 6, with increasing imaginary time, we can distinguish the peaks that are close to each other, aligning with our theoretical analysis.

In the main text, both the imaginary time $\tau$ and the cutoff $x_m$ are consistently adjusted by the target precision. We elaborate how to determine the the imaginary time $\tau$ and the cutoff $x_m$ by taking the Gaussian function as an example. For a certain simulation accuracy $\varepsilon$, the imaginary time $\tau$ and the cutoff $x_m$ satisfy $\tau \propto \sqrt{\log(2\Delta/\varepsilon)}$ and $x_m = 2\sqrt{\log(2/\varepsilon)}$, respectively. The reference line in the main text shows the error due to finite time $\tau$ plus the theoretical maximum error $L(\varepsilon)$.

We also show the simulation accuracy under the cooling process. Similarly in the main text, suppose we aim to find the second eigenstate $|u_2\rangle$, which has the largest overlap with the initial state. Given the associated eigenenergy $E_2$ found by the above method, we now analyze the simulation error during the cooling process. Here, the observable that verifies the state infidelity as $O = |u_2\rangle \langle u_2|$. We track the simulation accuracy by comparing with the ideal expectation value under the generalized cooling process, which is given by $\langle \psi_0 | g(\tau(H - E_2))\dagger O g(\tau(H - E_2)) | \psi_0 \rangle$. We can see that the error is kept in a relatively small value during the cooling process. We can improve the simulation accuracy by increasing the Monte Carlo sampling of the integral.