Measurement-based deterministic imaginary time evolution

Yuping Mao,1,2 Manish Chaudhary,1,2 Manikandan Kondapan,1,2 Junheng Shi,3,2 Ebubechukwu O. Ilo-Okeke,2,4 Valentin Ivanivikov,2,4 and Tim Byrnes2,1,4,5,6,7,∗

1State Key Laboratory of Precision Spectroscopy, School of Physical and Material Sciences, East China Normal University, Shanghai 200062, China
2New York University Shanghai, 567 West Yangsi Road, Shanghai, 200126, China
3CAS Key Laboratory of Theoretical Physics and Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China
4NYU-ECNU Institute of Physics at NYU Shanghai, 3663 Zhongshan Road North, Shanghai 200062, China
5Shanghai Frontiers Science Center of Artificial Intelligence and Deep Learning, 567 West Yangsi Road, Shanghai, 200126, China
6Center for Quantum and Topological Systems (CQTS), NYUAD Research Institute, New York University Abu Dhabi, UAE
7Department of Physics, New York University, New York, NY, 10003, USA

We introduce a method to perform imaginary time evolution in a controllable quantum system using measurements and conditional unitary operations. By performing a sequence of weak measurements based on the desired Hamiltonian constructed by a Suzuki-Trotter decomposition, an evolution approximating imaginary time evolution can be realized. The randomness due to measurement is corrected using conditional unitary operations, making the evolution deterministic. Both the measurements required for the algorithm and the conditional unitary operations can be constructed efficiently. We show that the algorithm converges only below a specified energy threshold and the complexity is estimated for some specific problem instances.

Introduction

Imaginary time evolution is an important and enduring concept in several areas of quantum physics, despite not being directly a physical process. In imaginary time evolution (ITE) of a quantum system with Hamiltonian $H$, time $t$ is replaced by imaginary time $t \to -i\tau$, such that the evolution operator is $e^{-iH\tau}$. As such, for long evolution times, the state approaches the ground state of the Hamiltonian. ITE can be directly applied as a numerical procedure on classical computers to obtain low-energy states. It is also central in making a formal connection between a $d$-spatial dimensional quantum field theory and a $d+1$-dimensional classical statistical mechanics system, through the Wick rotation. A variety of classical simulation methods take advantage of this connection, such as quantum Monte Carlo and its variants.

As a numerical procedure on a classical computer, ITE requires exponential resources that scale with the size of the Hilbert space. If there was a way of implementing ITE on a quantum computer efficiently, this could potentially be an extremely powerful tool. A direct implementation of the ITE operator $e^{-iH\tau}$, assuming elementary ITE gates, would have a complexity that scales polynomially with the number of subsystems, e.g. qubits. In comparison to the same calculation performed on a classical computer, this would give an exponential speedup. In ITE, convergence to a high fidelity state takes a timescale of the inverse energy gap. In a quantum simulation scenario, one is often interested in obtaining low-energy eigenstates of various systems, applicable to condensed matter physics, high-energy physics, and quantum chemistry. More generally, it may also be used as a general optimization tool, where a cost function is minimized. Applied to the context of solving the generalized Ising model, a problem that can be mapped to any optimization problem in the complexity class NP in polynomial time, the approach could be used to optimize problems in a variety of contexts such as logistics, financial applications, artificial intelligence, pharmaceutical and material development. Another application of ITE is as a state preparation protocol. For applications such as quantum metrology and alternative model of quantum computation, resource states need to be generated, which are sometimes difficult to produce. By engineering a suitable Hamiltonian where the desired state is the ground state, ITE can be used to generate and stabilize the state.

Several methods have been proposed to perform ITE in a controllable quantum system. In Variational Imaginary Time Evolution (VITE), McArdle,Yuan and co-workers introduced a hybrid quantum-classical approach to achieve ITE. Here, the Schrödinger equation is first solved in imaginary time on a classical computer to determine the parameters of a trial state, then this is used as the approximation of the quantum state for the quantum circuit. This method has been used to simulate the spectra of Hamiltonian, perform generalized time evolution, and to solve quantum many-body problems. Motta, Chan and co-workers proposed the Quantum Imaginary Time Evolution (QITE) method, where non-unitary time evolution is approximated by a unitary operator which contains the variation of the quantum systems. This method has been applied to the study of quantum simulation, nuclear energy level com-
computation [47], and quantum chemistry [48]. In another approach, Williams proposed a probabilistic approach to non-unitary quantum computing [53]. For example, in Probabilistic Imaginary Time Evolution (PITE) [54], an \( L \) qubit non-unitary gate simulation can be probabilistically obtained by designing an \( L + 1 \) qubit system and measuring the ancilla qubit \([55]\). When measuring the ancilla qubit, the \( L \)-qubit state will collapse into the desired state with a certain probability. PITE exploits Grover’s algorithm [56] to enhance the probability of getting the desired state while maintaining a high fidelity. PITE is suggested to be applicable to quantum chemistry problems [57]. The above ITE methods can be applied to various quantum algorithms. It has been shown that VITE can be applied to variational quantum algorithms for Boltzmann machine learning [58], while QITE can be applied to the QLanczos algorithm [40, 59] and variational quantum algorithms for Hamiltonian diagonalization [60].

In this paper, we propose a general method of performing ITE in a controllable quantum system. Our method relies upon performing measurements that mimic the ITE operator for small times. By performing repeated measurements on the system using these measurement operators, combined with a unitary correction step that acts conditionally on the measurement outcomes, this allows for a way to drive the state towards the lowest energy state of the given Hamiltonian. Much like quantum feed-forward approaches such as in quantum teleportation, this converts the stochastic evolution into a deterministic one, such that the desired state is obtained with unit probability for sufficiently long evolution times [61, 62]. The basic idea of the approach is to perform a weak measurement in the energy eigenbasis of a given Hamiltonian. During the slow collapse of the state, if the energy estimate is higher than a given threshold, then a conditional unitary is applied to disturb the system. This is repeated until the energy is sufficiently low, after which full collapse to the ground state occurs. Similar approaches were used for quantum state preparation [64] using weak measurements [65, 66]. Our approach differs from related works such as Refs. [54], where the desired outcome is obtained by postselection. It also differs from approaches such as in Refs. [3, 16] since the use of measurements involves an explicitly non-unitary step. As such, no precomputation needs to be performed to determine the evolution path.

**Weak energy measurements** We start by describing the general approach to performing ITE, then illustrate our approach with several examples. Our aim will be to perform ITE of an arbitrary Hamiltonian \( H \), such that we obtain the ground state
\[
e^{-H\tau}\left|\psi_0\right> \xrightarrow{\tau \to \infty} \left|E_0\right>,
\]
where \( \left|\psi_0\right> \) is an arbitrary initial state and \( \left|E_0\right> \) is the ground state of \( H \). We start by constructing measurement operators that take a similar form to the exponentiated Hamiltonian [11]. This can be achieved by performing a weak measurement of the Hamiltonian, with measurement operators
\[
M_0 = \langle 0\rangle_a e^{-ieH\otimes Y}|+\rangle_a = \frac{1}{\sqrt{2}}(\cos \epsilon H - \sin \epsilon H)
\]
\[
= \frac{1}{\sqrt{2}} \sum_n (\cos \epsilon E_n - \sin \epsilon E_n)|E_n\rangle\langle E_n| \approx e^{-\epsilon H}/\sqrt{2} \quad (2)
\]
\[
M_1 = \langle 1\rangle_a e^{-ieH\otimes Y}|+\rangle_a = \frac{1}{\sqrt{2}}(\cos \epsilon H + \sin \epsilon H)
\]
\[
= \frac{1}{\sqrt{2}} \sum_n (\cos \epsilon E_n + \sin \epsilon E_n)|E_n\rangle\langle E_n| \approx e^{\epsilon H}/\sqrt{2}. \quad (3)
\]
where Pauli spin operators are denoted \( X, Y, Z \), and the approximation is valid for \( |\epsilon H| \ll 1 \). The Hamiltonian is taken to have a suitable energy offset and \( \epsilon \) is chosen such that the energy spectrum fits in the region \(-\pi/4 \leq \epsilon E_n \leq \pi/4\). This measurement can be realized by preparing an ancilla qubit in the state \( |+\rangle_a = (|0\rangle_a + |1\rangle_a)/\sqrt{2} \) and performing an interaction with Hamiltonian \( H \otimes Y \), and measuring the ancilla in the \( Z \)-basis. The measurement operators satisfy \( M_0^\dagger M_0 + M_1^\dagger M_1 = I \), where \( I \) is the identity matrix. In the case that the interaction \( H \otimes Y \) is not directly accessible due to the Hamiltonian being composed of a sum of terms \( H = \sum_{j=1}^{N} H(j) \), a Suzuki-Trotter decomposition [67, 68] of \( e^{-i\epsilon H(j)\otimes Y} \) to suitable order is instead performed (see Supplementary Information). This has the effect of changing the precise form of (3), but is still an approximation to the imaginary time exponentiated Hamiltonian.

We wish to perform the ITE to amplify the ground state as in (1). If it were possible to apply \( M_0 \) only, this would achieve a similar evolution to (1) since \( \cos \epsilon E_n - \sin \epsilon E_n \) is monotonically decreasing in the domain \(-\pi/4 \leq \epsilon E_n \leq \pi/4\), such that \( M_0^k|\psi_0\rangle \xrightarrow{k \to \infty} |E_0\rangle \). However, since the two outcomes \( \{M_0, M_1\} \) occur randomly according to quantum measurement probabilities,

![FIG. 1](image-url)
such a sequence is typically a rare occurrence. Let us analyze a particular measurement sequence where there are $k_0$ counts of $M_0$ and $k_1$ counts of $M_1$. Since $[M_0, M_1] = 0$, the order of the outcomes does not matter and this measurement sequence can be written

$$M_0^{k_0} M_1^{k_1} |\psi_0\rangle = \sum_n A_{k_0 k_1} (\epsilon E_n) |E_n| |\psi_0\rangle |E_n\rangle \quad (4)$$

where we defined the amplitude function

$$A_{k_0 k_1}(x) = \frac{1}{\sqrt{2k_0+k_1}} (\cos x - \sin x)^{k_0} (\cos x + \sin x)^{k_1}
= \cos^{k_0}(x + \pi/4) \sin^{k_1}(x + \pi/4). \quad (5)$$

In Fig. 1(a) we show a plot of the function $A$. We see that for $-\pi/4 \leq x \leq \pi/4$ and a large number of measurements it has a Gaussian form [69], where the peak value occurs at $x_{k_0k_1}^{\max} = \epsilon E_{k_0k_1}^{\max} = \frac{1}{2} \arcsin \left( \frac{k_1 - k_0}{k_0 + k_1} \right) \quad (6)$

and the width is $\sigma = 1/\sqrt{2(k_0 + k_1)}$. Here, $E_{k_0k_1}^{\max}$ is the peak value in terms of energy. As the number of measurements are increased, the Gaussians become increasingly well-defined (Fig. 1(b)). In the limit of a large number of measurements, a collapse on the energy basis occurs.

In order to increase the amplitude of the ground state in (4), we require that the Gaussian is peaked with an outcome with $x_{k_0k_1}^{\max} < \epsilon(E_0 + E_1)/2$ (see Fig. 1(b)). This will create an amplitude gain of the ground state over all the remaining states, since the peak of Gaussian is closer to $\epsilon E_0$ than any other eigenvalue, and the tail of the Gaussian on the higher energy side will suppress all higher energy states. So our strategy will then be to control the position of the Gaussian such that it lies in the desired energy range.

The algorithm To this end, we turn to an adaptive strategy, where a unitary operation is applied conditioned on the measurement outcomes. Our basic strategy will be to continually monitor the location of the Gaussian using the expression (6). If the location of Gaussian corresponds to a sufficiently low energy state, then no unitary is applied. If the Gaussian is located at a value that is of a higher energy than a chosen energy threshold $E_{\text{th}}$, then a corrective unitary is applied. Concretely, we iteratively perform

$$|\psi_{t+1}\rangle = \frac{U_{k_0k_1(t+1)}^{(t+1)} M_n |\psi_t\rangle}{\sqrt{\langle \psi_t | M_n^\dagger M_n | \psi_t \rangle}} \quad (7)$$

where $n \in \{0, 1\}$ labels the $(t+1)$th measurement outcome, with

$$U_{k_0k_1} = \begin{cases} I & \text{if } x_{k_0k_1}^{\max} < \epsilon E_{\text{th}} \\ U_C & \text{otherwise} \end{cases} \quad (8)$$

and

$$k_m^{(t+1)} = \begin{cases} k_m^{(t)} + \delta_{m}^{(t)} & \text{if } x_{k_0k_1}^{\max}< \epsilon E_{\text{th}} \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

are the cumulative measurement outcomes starting with $k_m^{(0)} = 0$. In words, this counts the number of $M_0, M_1$ measurements respectively, until it is found that the energy estimate is above the threshold, at which point the counts are reset to zero. To ensure convergence of the sequence to the ground state, we demand a non-zero transition amplitude between all energy eigenstates $|\{E_n |U_C| E_m\rangle| > 0, \forall n, m$. For $E_0 < E_{\text{th}} < E_1$, this ensures that only the ground state is the unique fixed point of the evolution (see Supplementary Material). The requirement $|\{E_n |U_C| E_m\rangle| > 0$ is not usually very difficult to satisfy since it merely requires off-diagonal matrix elements in the energy basis, which occurs for a large number of matrices. Practically, one may choose a random unitary matrix based on readily available gates. In this way, the wavefunction for the ground state does not need to be known for the procedure. We note that if there is some knowledge of the eigenstates $|E_n\rangle$, then more sophisticated strategies beyond the above requirement and (8) can be used to construct $U_C$. For instance, rotations targeting the ground state based on the energy estimate $E_{k_0k_1}^{\max}$ could be implemented.

Example 1: One qubit We start with the simplest example of a single qubit with Hamiltonian $H = Z$. In Fig. 2(a)(b) we show the evolution of the states on the Bloch sphere for the measurements $M_0, M_1$. We see that $M_0$ has the effect of driving all states towards the south pole of the Bloch sphere, while $M_1$ drives all states to the north pole, following longitudinal lines. This is consistent with the imaginary time operator $e^{tZ}$, as given in (7) and (8). In Fig. 2(c), we show the fidelity $F = |\langle \psi_t | E_0 \rangle|^2$ for three different measurement sequences. Due to the randomness of quantum measurements, each sequence gives a different trajectory, but all cases converge to the ground state $|E_0\rangle = |1\rangle$. Averaging over many random trajectories yields a smooth exponential curve approaching the target state. A semilog plot (Fig. 2(c) inset) verifies the exponential evolution, consistent with ITE. In fact, for this case it can be shown exactly that any trajectory is equivalent to applying a power of $M_0$ which approaches the ground state (see Supplementary Material). In Fig. 2(d) we plot the peak position of the $A$ function for the same three trajectories as in Fig. 2(c). We see that there are broadly two regimes where there is a random movement of the peak position, followed by a region of stability, where the Gaussian approaches the ground state energy. In the initial random evolution, when $x_{k_0k_1}^{\max} > \epsilon E_{\text{th}}$, several spin flips induced by $U_C$ occur, until the random movement stabilizes to the correct energy range. After the correct peak position is estab-
Established, the fidelity quickly evolves towards the ground state.

**Example 2: Transverse-field Ising model** We next show an example of the transverse-field Ising model with the Hamiltonian $H^{(1)} = \lambda \sum_{n=1}^{N} X_n$, $H^{(2)} = \sum_{n=1}^{N-1} Z_n Z_{n+1}$, $H = H^{(1)} + H^{(2)}$. Here, $L$ is the number of qubits in the chain, and we take $E_{th} = (E_0 + E_1)/2$. We assume that each of the terms in the Hamiltonian must be implemented separately to construct the measurement operators. We perform a second order Suzuki-Trotter expansion with $M_n = \langle n | e^{-i\epsilon H^{(2)}} \otimes Y/2 | e^{-i\epsilon H^{(1)}} \otimes Y/2 | n \rangle$ (see Supplementary Materials). The conditional operator is chosen to be a random local unitary $U_C = \sum_{n=1}^{N} e^{i \frac{2\pi}{\epsilon n} (\phi_n X_n + \phi_n Y_n + \phi_n Z_n)}$, where $\phi_n \in [0, 1]$. We show the fidelity of the procedure with respect to the target state in Fig. 3(a). Again we see two stages where there is a random evolution of the fidelity, followed by a smoother time evolution once the peak of Gaussian amplitude function is in the correct range. For longer chains we observe a longer period of random evolution before the correct energy range is established, after which the system quickly converges to the ground state.

**Complexity estimate** We now briefly discuss the complexity of the proposed algorithm. First, the measurements $M_n$ can be typically performed efficiently for a given Hamiltonian using a Suzuki-Trotter decomposition (see Supplementary Information). Due to the flexibility of the choice of the operator $U_C$, this can also be typically be implemented efficiently. The complexity of the algorithm then results from the number of measurements that need to be made in total. Based on the behavior observed in Figs. 2 and 3 we model the initial part of the measurement sequence as a stochastic process, where the algorithm repeats until the criterion $x_{k_1}^{max} < \epsilon E_{th}$ is satisfied (see Supplementary Material). The number of required measurements until this occurs can be estimated by evaluating the probability of obtaining a sequence with $k$ consecutive $M_0$ outcomes together with the average failed sequence length. Although it is not easy to obtain a simple expression for the general case complexity, for two particular cases, assuming an initial state with equal superposition, it is possible to estimate the typical number of measurements before convergence. These are Hamiltonians with (I) a uniform density of states and (II) a completely degenerate spectrum of excited states (see Supplementary Information). For (I), we obtain a scaling as $O(1/(\epsilon \Delta)^2)$, where $\Delta = E_1 - E_0$ is the gap. We note that there is an implicit dependence upon system dimension in this relation, due to the requirement that $-\pi/4 \leq \Delta E_n \leq \pi/4$. For example, for an exponential number of states, $\epsilon \Delta$ is exponentially vanishing and the final scaling increases exponentially for unstructured problems. For (II), we find that the scaling is $O(D)$, where $D$ is the system dimension.

**Conclusions** We have proposed a method of performing deterministic ITE, using measurements and conditional unitary operations. Due to use of quantum measurements, the evolution is stochastic within Hilbert space on a shot-to-shot basis. Averaging over trajectories reveals an exponential evolution that is consistent with ITE. The approach is generic, one does not need to know the ground state before executing the algorithm, and the measurement operators can be constructed with...
a Suzuki-Trotter decomposition so that it is compatible with gate based quantum computing. The measurement operators and unitary operators can be constructed efficiently, but the number of measurements that need to be performed before convergence depends upon the nature of the Hamiltonian and the initial state. The algorithm is guaranteed to only converge if the energy of the state is lower than $E_{th}$.

The algorithm that we present here can be considered a generalization of several related works which use the same basic framework. For example in Ref. [7], a similar method was proposed to generate supersinglet states, and also maximally entangled states of atomic ensembles in Ref. [10]. A four-qubit linear graph state was also deterministically generated using the method in Ref. [7].

We have found that the algorithm converges to the ground state for every problem Hamiltonian that we have given it. In our algorithm, we chose a relatively simple strategy for the adaptive unitary operator $\hat{U}$, where the state is rotated if the measurement outcomes do not fall in the targeted range. Since $E_{\text{th}}^{\max}$ is an energy estimate of the state, more complex strategies to rotate the state to the ground state could be made. Another potential improvement is to choose a judicious initial state to improve the convergence of the scheme.

This work is supported by the National Natural Science Foundation of China (62071301); NYU-ECNU Institute of Physics at NYU Shanghai; Shanghai Frontiers Science Center of Artificial Intelligence and Deep Learning; the Joint Physics Research Institute Challenge Grant; the Science and Technology Commission of Shanghai Municipality (19XD1423000,22ZR1444600); the NYU Shanghai Boost Fund; the NYU Shanghai Major-Grants Seed Fund; Tamkeen under the NYU Abu Dhabi Research Institute grant CG008; and the SMEC Scientific Research Innovation Project (2023ZKZD55). J.S. is supported by the National Natural Science Foundation of China Grant Nos. 11925507 and 12047503.

\textsuperscript{*} tim.byrnes@nyu.edu

[1] J. J. Sakurai and E. D. Commins, 	extit{Modern quantum mechanics, revised edition} (1995).
[2] W. Magnus, Communications on pure and applied mathematics 7, 649 (1954).
[3] G. Vidal, Physical review letters 98, 070201 (2007).
[4] S.-H. Lin, R. Dilip, A. G. Green, A. Smith, and F. Pollmann, PRX Quantum 2, 010342 (2021).
[5] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Physical review letters 98, 140506 (2007).
[6] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, npj Quantum Information 5, 1 (2019).
[7] M. L. Chiofalo, S. Suzuki, and M. Tosi, Physical Review E 62, 7438 (2000).
[8] S. Palpacelli, S. Suzuki, and R. Spigler, Phys. Rev. E 76, 036712 (2007).
[9] J. Liu and N. Makri, Molecular Physics 103, 1083 (2005).
[10] G.-C. Wick, Physical Review 96, 1124 (1954).
[11] M. E. Peskin, 	extit{An introduction to quantum field theory} (CRC press, 2018).


[39] M. Tame, M. Paternostro, M. Kim, and V. Vedral, Physical Review A 73, 022309 (2006).
[40] S. D. Bartlett and T. Rudolph, Physical Review A 74, 040302 (2006).
[41] M. Van den Nest, K. Luttmmer, W. Dür, and H. Briegel, Physical Review A 77, 012301 (2008).
[42] T. H. Kyaw, Y. Li, and L.-C. Kwek, Physical Review Letters 113, 180501 (2014).
[43] T. Jones, S. Endo, S. McArdle, X. Yuan, and S. C. Benjamin, Physical Review A 99, 062304 (2019).
[44] S. Endo, J. Sun, Y. Li, S. C. Benjamin, and X. Yuan, Physical Review Letters 125, 010501 (2020).
[45] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, Quantum 3, 191 (2019).
[46] M. Motta, C. Sun, A. T. Tan, M. J. O’Rourke, E. Ye, A. J. Mimmich, F. G. Brudnadö, and G. K.-L. Chan, Nature Physics 16, 205 (2020).
[47] K. Yeter-Aydéniz, R. C. Pooser, and G. Siopsis,npj Quantum Information 6, 1 (2020).
[48] N. Gomes, F. Zhang, N. F. Berthusen, C.-Z. Wang, K.-M. T. H. Kyaw, Y. Li, and L.-C. Kwek, Physical Review Letters 125, 010501 (2020).
[49] K. Yeter-Aydéniz, G. Siopsis, and R. C. Pooser, New Journal of Quantum Information II (International Society for Optics and Photonics, 2004), vol. 5436, pp. 297–306.
[50] L. K. Grover, Phys. Rev. Lett. 79, 325 (1997).
[51] T. Kosugi, Y. Nishiya, and Y.-i. Matsushita, arXiv preprint [arXiv:2111.12470] (2021).
[52] Y. Shingu, Y. Seki, S. Watabe, S. Endo, Y. Matsuzaki, T. Kosugi, E. Anisimova, et al., Nature 500, 319 (2013).
[53] E. O. Ilo-Okeke, L. Tessler, J. P. Dowling, and T. Byrnes, npj Quantum Information 4, 1 (2018).
[54] E. O. Ilo-Okeke and T. Byrnes, Physical Review letters 112, 233602 (2014).
[55] E. O. Ilo-Okeke and T. Byrnes, Physical Review A 94, 013617 (2016).
[56] M. Suzuki, Physics Letters A 180, 232 (1993).
[57] E. Kapit, P. Ginsparg, and E. Mueller, Physical Review Letters 108, 066802 (2012).
[58] M. A. Ochoa, W. Belzig, and A. Nitzan, Scientific reports 8, 1 (2018).
[59] E. O. Ilo-Okeke, P. Chen, S. Li, B. C. Anusionwu, V. Ivannikov, and T. Byrnes, AVS Quantum Science 5 (2023).
[60] M. Chaudhary, E. O. Ilo-Okeke, V. Ivannikov, and T. Byrnes, arXiv preprint [arXiv:2302.07526] (2023).
[61] M. Kondappan, M. Chaudhary, E. O. Ilo-Okeke, V. Ivannikov, and T. Byrnes, Physical Review A 107, 042616 (2023).

SUPPLEMENTARY INFORMATION

CONSTRUCTION OF MEASUREMENT OPERATORS

Derivation of Eqs. (2) and (3)

To realize the measurement operator in Eqs. (2) and (3) of the main text, we use an ancilla qubit to perform a weak measurement of the energy via a Hamiltonian of the form $H \otimes Y$. The ancilla qubit is initially prepared in the state $|+\rangle_a = (|0\rangle_a + |1\rangle_a)/\sqrt{2}$, where $Y$ is the Pauli-Y matrix. Interacting the total Hamiltonian for a time $\epsilon$ we have

\begin{equation}
\begin{split}
e^{-iH\otimes Y\epsilon}|\psi_0\rangle|+\rangle_a &= \sum_n e^{-iE_n\epsilon Y} \langle E_n|\psi_0\rangle|E_n\rangle|+\rangle_a \\
&= \frac{1}{\sqrt{2}} \sum_n \langle E_n|\psi_0\rangle|E_n\rangle \left[ (\cos E_n\epsilon - \sin E_n\epsilon)|0\rangle_a + (\cos E_n\epsilon + \sin E_n\epsilon)|1\rangle_a \right].
\end{split}
\end{equation}

(S1)

Projecting on the ancilla onto the $|0\rangle_a$ state gives

\begin{equation}
\begin{split}
|0\rangle_a \langle 0| e^{-iH\otimes Y\epsilon}|\psi_0\rangle|+\rangle_a \\
&= \frac{1}{\sqrt{2}} \sum_n \langle E_n|\psi_0\rangle(\cos E_n\epsilon - \sin E_n\epsilon)|E_n\rangle|0\rangle_a
\end{split}
\end{equation}

(S2)

while the $|1\rangle_a$ outcome gives

\begin{equation}
\begin{split}
|1\rangle_a \langle 1| e^{-iH\otimes Y\epsilon}|\psi_0\rangle|+\rangle_a \\
&= \frac{1}{\sqrt{2}} \sum_n \langle E_n|\psi_0\rangle(\cos E_n\epsilon + \sin E_n\epsilon)|E_n\rangle|1\rangle_a.
\end{split}
\end{equation}

(S3)

The ancilla qubit decouples from the system after the measurement and the combined effect can be given according to the expressions given in Eq. (2) and (3) of the main text.

Suzuki-Trotter decomposition

The approach of the previous section can be used to construct the measurement operator if the Hamiltonian $H \otimes Y$ is readily implementable. For complex Hamiltonians involving many terms $H = \sum_n H^{(n)}$, it may be necessary to construct the total Hamiltonian evolution via
a Suzuki-Trotter decomposition. We show in this section the measurement operators for this case.

For a Hamiltonian consisting of two non-commuting terms we may perform a first order Suzuki-Trotter decomposition to give the measurement operator

\[
M_n = \langle n | e^{-i H_0^{(1)} \otimes Y / 2} e^{-i H_0^{(2)} \otimes Y} e^{-i H_0^{(3)} \otimes Y / 2} | + \rangle_a
\]

\[
= \frac{1}{\sqrt{2}} \left( \cos \frac{\epsilon H^{(1)}}{2} \cos \frac{\epsilon H^{(2)}}{2} - \frac{\epsilon H^{(1)}}{2} \sin \frac{\epsilon H^{(2)}}{2} \right)
\]

\[
- (-1)^n \frac{1}{\sqrt{2}} \left( \cos \frac{\epsilon H^{(1)}}{2} \sin \frac{\epsilon H^{(2)}}{2} + \frac{\epsilon H^{(1)}}{2} \cos \frac{\epsilon H^{(2)}}{2} \right)
\]

for \( n \in \{0, 1\} \). This has errors at the level of \( O(\epsilon^3) \). The superior convergence of the second order Suzuki-Trotter form is attributed to the fact that \( M_0 \) and \( M_1 \) have the same eigenstates for \( S_3 \) but not \( S_2 \), resulting in better stability of the fixed point.

\[
\text{GROUND STATE AS THE UNIQUE FIXED POINT OF EVOLUTION}
\]

Assume that \( E_0 < E_{\text{th}} < E_1 \), where \( E_0 \) is the ground state energy and \( E_1 \) is an excited state that is not degenerate with the ground state, i.e. \( E_1 > E_0 \). First limiting ourselves to energy eigenstates \( |E_n\rangle \), we show that only the ground state \( |E_0\rangle \) (and its degenerate states) are fixed points of the iteration. We then later generalize to more general states and show that only energy eigenstates need to be considered.

First consider the case of the initial state being the ground state \( |\psi_0\rangle = |E_0\rangle \). After \( K = k_0 + k_1 \) applications of the measurement operator one obtains

\[
U_{k_0 k_1} M_0^{k_0} M_1^{k_1} |E_0\rangle = A_{k_0 k_1} (\epsilon E_0) U_{k_0 k_1} |E_0\rangle,
\]

where we used Eq. (4) in the main text. To see the effect of the unitary \( U_{k_0 k_1} \), consider the most likely outcome of the measurement sequence \( k_0 \approx p_0 K, k_1 \approx p_1 K \) where \( K \) is the total number of measurements. Here, the probabilities of obtaining the two measurement outcomes for various eigenstates is

\[
p_0 = \langle E_n | M_0^\dagger M_0 | E_n \rangle = \frac{1}{2} \left( 1 - \sin 2 \epsilon E_n \right)
\]

\[
p_1 = \langle E_n | M_1^\dagger M_1 | E_n \rangle = \frac{1}{2} \left( 1 + \sin 2 \epsilon E_n \right).
\]

Substituting these values into Eq. (6) of the main text, this gives

\[
x_{k_0 k_1}^{\text{max}} = \frac{1}{2} \arcsin \left( \frac{k_1 - k_0}{k_0 + k_1} \right)
\]

\[
= \frac{1}{2} \arcsin \left( p_1 - p_0 \right)
\]

\[
= \epsilon E_0.
\]

We see that this obeys \( x_{k_0 k_1}^{\text{max}} = \epsilon E_0 < \epsilon E_{\text{th}} \) such that \( U_{k_0 k_1} = I \) according to Eq. (8) of the main text. Therefore in this case

\[
U_{k_0 k_1} M_0^{k_0} M_1^{k_1} |E_0\rangle \propto |E_0\rangle,
\]

and the ground state is a fixed point of the evolution.

For any initial state that is an excited state \( |E_n\rangle \) with \( n > 0 \), using similar arguments to \( S_3 \), the measurement readouts converge to \( x_{k_0 k_1}^{\text{max}} = \epsilon E_n \). However, if \( x_{k_0 k_1}^{\text{max}} \geq \epsilon E_{\text{th}} \), the state is rotated away from \( |E_n\rangle \), since by definition \( \langle (E_n | M_C | E_m) \rangle > 0 \), \( \forall n, m \), and the measurement operators are diagonal in the energy basis. Hence any excited state is not a fixed point of the iteration.

Now consider the more general case of an arbitrary state. Since the measurement operators are diagonal in the energy basis, for the case that \( x_{k_0 k_1}^{\text{max}} < \epsilon E_{\text{th}} \) where \( U_{k_0 k_1} = I \), the ground state is the only energy eigenstate which is a fixed point. Then the only possibility is that the state \( |\psi\rangle \) is an eigenstate of the combination of \( U_{C} \) and the measurement operators

\[
U_C M_0 |\psi\rangle \propto |\psi\rangle
\]

\[
U_C M_1 |\psi\rangle \propto |\psi\rangle,
\]

where we must consider both possibilities since in general either outcome may occur. Such a state must produce the same state (up to a proportionality factor) for either
measurement operator. For an arbitrary state we may evaluate
\[ U_CM_0|\psi \rangle = \sum_n (E_n|\psi \rangle \langle \cos \epsilon E_n - \sin \epsilon E_n|C_n \rangle \]
\[ U_CM_1|\psi \rangle = \sum_n (E_n|\psi \rangle \langle \cos \epsilon E_n + \sin \epsilon E_n|C_n \rangle \]
where \(|C_n \rangle = U_C|E_n \rangle\) are a set of orthogonal basis states. It is only possible to have \(U_CM_0|\psi \rangle \propto U_CM_1|\psi \rangle\) if \(|\psi \rangle\) is one of the energy eigenstates, which removes the \(n\) dependence and hence the factor of \(\cos \epsilon E_n + \sin \epsilon E_n\) after normalization. Since we have already shown that the only energy eigenstate that is a fixed point of the evolution is \(|E_0 \rangle\), this completes the proof.

**COMPLEXITY ESTIMATE OF THE ALGORITHM**

We now estimate the complexity of the algorithm for two prototypical Hamiltonian spectra. Firstly, the measurement operators \(M_n\) can be constructed efficiently using the methods given in the first section of supplementary information. Even in the case that a Suzuki-Trotter decomposition is used to construct the measurement operators, this will scale polynomially with the number of terms in the Hamiltonian, which typically scales polynomially with the number of qubits. In this way the measurement operators can be constructed efficiently. The \(U_C\) operator can also be chosen according to what gates are available such that it satisfies \(|\langle E_n|U_C|E_m \rangle| > 0, \forall n, m\). Thus the unitary correction operator can also be chosen to be implemented efficiently with the number of qubits.

The main complexity of the algorithm is then determined by the number of iterations required before convergence is attained. A hint of the dynamics towards convergence can be seen in Figs. 2 and 3 of the main text. We see here that initially there is a period of chaotic evolution and the peak position. At some number of iterations, the measurement sequence “locks in”, after which rapid convergence towards a fidelity of 1 is attained.

We can understand the dynamics in the following way. We wish to obtain imaginary time evolution by taking advantage of the similarity of \(M_0\) with the operator \(e^{-\epsilon tH}\). However, since measurements are random, we are not guaranteed to obtain only \(M_0\) and one will obtain \(M_1\) with some probability. The algorithm as given in Eqns. (7)-(9) of the main text keeps trying until one obtains a sequence satisfying \(x_{k_0k_1}^{\text{max}} < \epsilon E_{\text{th}}\), which amounts to convergence of a low energy state below a particular threshold. The algorithm keeps iterating until one obtains the desired sequence where there is a sufficiently large number of \(M_0\) measurements over \(M_1\) measurements.

For a state that has a uniform amplitude in energy eigenstates (i.e. \(|\langle \psi |E_n \rangle|\) = constant), the probability of obtaining \(M_0\) and \(M_1\) is \(1/2\). This might naively suggest that to obtain \(k\) measurements that are all \(M_0\) has a probability of \(1/2^k\), which is vanishing for a large number of measurements. This is however a large underestimate, due to the way quantum mechanical measurements work. In fact, getting the \(M_0\) outcome is a self-enhancing process, where each time an \(M_0\) outcome occurs, it becomes more likely to obtain \(M_0\) again.

**Uniform density of states**

To see this quantitatively, let us consider the following example. We choose a problem in a \(D\) dimensional Hilbert space with uniformly distributed energies. The Hamiltonian is
\[ H = \sum_{n=0}^{D-1} E_n |n\rangle \langle n| \]
where for a uniform distribution we take
\[ E_n = 2n - D + 1 \]
such that the ground state \(E_0 = -(D - 1)\) and the highest energy state is \(E_{D-1} = D - 1\). We may choose \(\epsilon\) such that the spectrum occupies the full range of the domain of \(A_{k_0k_1}(\epsilon)\), which is \(-\pi/4 \leq \epsilon \leq \pi/4\). Namely, \(\epsilon = \pi/(4(D - 1))\) such that \(\epsilon E_0 = -\pi/4\) and \(\epsilon E_{D-1} = \pi/4\). This maximizes \(\epsilon\) and typically gives the best performance of the algorithm. As discussed in the main text, we require choosing the energy threshold \(E_{\text{th}} < (E_0 + E_1)/2\) such that the ground state has the largest amplification factor. With this choice, under the Gaussian approximation of the function \(A_{k_0k_1}(\epsilon)\), the amplitude function obeys \(A_{k_0k_1}(\epsilon E_0) > A_{k_0k_1}(\epsilon E_n)\) for \(n > 0\) and the ground state is amplified according to Eq. (4) of the main text. We consider a case with \(D \gg 1\) and a uniform density of states, such that we may take \(E_{\text{th}} \approx E_0\), which means that the condition in Eq. (8) of the main text is \(x_{k_0k_1}^{\text{max}} < -\pi/4\). Translating this to \(k_0, k_1\) using Eq. (6) of the main text, this means that one must obtain every measurement outcome to be \(M_0\), and hence \(k_1 = 0\) to satisfy the convergence.

Let us calculate the probability of obtaining \(k\) measurement outcomes that are \(M_0\). For simplicity, suppose the initial state is in the state
\[ |\psi_0 \rangle = \frac{1}{\sqrt{D}} \sum_{n=0}^{D-1} |n\rangle. \]
Then after \(k\) outcomes for the measurement \(M_0\), the resulting state is
\[ M_0^k|\psi_0 \rangle = \frac{1}{\sqrt{D}} \sum_{n=0}^{D-1} \cos(k\epsilon E_n + \pi/4)|n\rangle. \]
The probability of this measurement outcome, where there are \( k \) consecutive \( M_0 \) outcomes, is
\[
p_k = \langle \psi_0 \vert (M_0^\dagger)^k M_0^k \vert \psi_0 \rangle
= \frac{1}{D} \sum_{n=0}^{D-1} \cos^{2k}(\epsilon E_n + \pi/4). \tag{S16}
\]
Let us assume a uniform density of states in energy space, such that the sum can be approximated by
\[
p_k \approx \frac{1}{\pi/2} \int_{-\pi/4}^{\pi/4} dx \cos^{2k}(x + \pi/4)
= \frac{1}{4^k} \binom{2k}{k}
\approx \frac{1}{\sqrt{k\pi}}. \tag{S17}
\]
where the last approximation is valid for \( k \gg 1 \). We observe that the probability of obtaining consecutive \( M_0 \) outcomes is in fact much larger than the \( 1/2^k \) estimate one would naively make from an independent probability assumption.

The self-enhancing effect of consecutive \( M_0 \) outcomes is a weak measurement version of the familiar effect known for projective measurements. For example, consider the projective measurements \( P_0 = \vert 0 \rangle \langle 0 \vert \) and \( P_1 = I - \vert 0 \rangle \langle 0 \vert \). If the measurement outcome \( P_0 \) occurs on an initial state \( \psi_0 \propto \vert 0 \rangle \), then with unit probability subsequent outcomes will all be \( P_0 \). In the weak measurement case that we consider, the enhancement of probability is more gradually attained, rather than after a single measurement.

Returning to the case that we consider here, any time the measurement outcome \( M_1 \) is obtained, we violate the criterion \( \epsilon_{max}^{k_{0k_1}} < \epsilon E_{th} = \epsilon E_0 = -\pi/4 \), and a new attempt at convergence begins. We may estimate the total number of measurements before a sequence \( M_0^k \) is obtained as \( 1/p_k \), multiplied by the average length of a failed sequence. A failed sequence consists of any sequence which is shorter than the target length and ends in an \( M_1 \) outcome. Specifically, if we consider \( M_0^k \) to be a successful sequence, then the sequences \( M_1, M_1 M_0, M_1 M_0^2, \ldots, M_1 M_0^{k-1} \) to be the associated failed sequences. To evaluate the probability of a failed measurement sequence, first evaluate the probability of obtaining \( M_0 \) following \( M_0^k \) as
\[
p_{0|k} = \frac{p_{k+1}}{p_k} = 1 - \frac{1}{2(k+1)}, \tag{S19}
\]
where we used (S17). The probability of obtaining \( M_1 \) following \( M_0^k \) is then
\[
p_{1|k} = 1 - p_{0|k} = \frac{1}{2(k+1)}. \tag{S20}
\]
The probability of obtaining an outcome \( M_1 M_0^{k-1} \) is then
\[
p_{1|k-1} p_{k-1} = \frac{2}{k^4} \left( \frac{2k - 2}{k - 1} \right). \tag{S21}
\]
The total failure probability for a sequence of length \( k \) is then
\[
p_{k}^{\text{fail}} = \sum_{k'=1}^{k} p_{1|k'-1} p_{k'-1}
= 1 - \frac{1}{4^k} \left( \frac{2k}{k} \right) \tag{S22}
\]
which is equal to \( 1 - p_k \) as expected. The average length of a failed sequence is then
\[
T_k^{\text{fail}} = \sum_{k'=1}^{k} k' p_{1|k'-1} p_{k'-1}
= \frac{k}{4^k} \left( \frac{2k}{k} \right)
\approx \sqrt{\frac{k}{\pi}} \tag{S23}
\]
The expected total number of measurements before a target sequence of \( M_0^k \) is obtained is then
\[
T = \frac{T_k^{\text{fail}}}{p_k} + k \approx 2k, \tag{S24}
\]
where we have added a \( k \) to the total to account for the number of measurements in \( M_0^k \) itself.

Now let us estimate what \( k \) should be such that the ground state is obtained with high fidelity. For large \( k \), we may approximate (S15) with a Gaussian such that
\[
M_0^k \vert \psi_0 \rangle = \frac{1}{\sqrt{D}} \sum_{n=0}^{D-1} \exp\left(-\frac{(\epsilon E_n + \pi/4)^2}{2\sigma^2}\right) \vert n \rangle. \tag{S25}
\]
where the standard deviation is \( \sigma = 1/\sqrt{k} \). To obtain a high fidelity convergence to the ground state, we require that at the energy of the first excited state \( E_1 \), the Gaussian sufficiently suppresses its amplitude. Hence we require
\[
\epsilon(E_1 - E_0) \sim 2\sigma = \frac{2}{\sqrt{k}}, \tag{S26}
\]
where we have put a \( 2\sigma \) standard deviation which suppresses the first excited state by a factor 0.14.

Putting together (S24) and (S26) we obtain the expected total number of measurements
\[
T \approx \frac{8}{(\epsilon \Delta)^2}, \tag{S27}
\]
where \( \Delta = E_1 - E_0 \).
Naively, the uniform density of states Hamiltonian is a constant gap problem with \( \Delta = 2 \), and is independent of \( D \), according to Eq. (S13). Hence it appears that the scaling of the imaginary time evolution allows for a way to solve the problem independent of problem dimension \( D \), which we consider to be an exponentially large quantity. However, an important point is that we must also choose an \( \epsilon \) such that the energy spectrum of the Hamiltonian lies in the range \(-\pi/4 \leq \epsilon E_n \leq \pi/4\). As obtained previously, here we require \( \epsilon \leq \pi/(4(D-1)) \). This means that even if the gap \( \Delta \) itself is a constant as the problem size \( D \) grows, we must fit an exponential number of states within the range \(-\pi/4 \leq \epsilon E_n \leq \pi/4\). Therefore the combination \( \epsilon \Delta \) is an exponentially small quantity for exponentially large \( D \). Then according to our scaling \( O(1/(\epsilon \Delta)^2) \) as written in the main text, the algorithm would take an exponential time for convergence.

We have performed some numerical analysis to verify the above scaling. In Fig. S1 we show the most probable number of measurements to first attain a fidelity of 0.9. To obtain this, we run the algorithm as described in the main text applied to the Hamiltonian (S12) many times and find what is the most likely number of measurements. In the case of Fig. S1(a), we fix \( D \) and vary \( \epsilon \) in the range \( 0 < \epsilon \leq 1/(\pi(D-1)) \). We see that the numerical values follow the predicted scaling \( T \propto 1/(\epsilon \Delta)^2 \). For Fig. S1(b), we take \( \epsilon = \pi/(4(D-1)) \) and vary \( D \). In terms of \( D \), we therefore expect \( T \propto (D-1)^2 \) from (S27). We see excellent agreement to the numerical data, again taking the most probable number of measurements to obtain a fidelity of 0.9.

The Hamiltonian (S12) does not appear to be a computationally difficult problem due to its rather simple energy structure. It can however be converted to a computationally difficult problem by reassigning the energies randomly to the states. In this case, the problem has no structure and would be a computationally hard problem. Specifically, the Hamiltonian reads

\[
H = \sum_{n=0}^{D-1} E_n |P(n)\rangle \langle P(n)|
\]  

where \( P(n) \) is a permuting function which rearranges the states in a randomized manner with no structure. The energy distribution is uniform as before, given in Eq. (S18). In this case formally the analysis is the same as above except that the energy labels are reordered. The analysis can be repeated in the same way, such that we obtain the same scaling. We obtain the same result since the simple energy structure was never exploited (in constructing \( U_C \) for example), hence the scaling remains the same.

We emphasize that the assumptions of the uniformly distributed energies and even superposition initial state are not required for the running of our algorithm itself.

Next we examine the opposite limit of a golf course Hamiltonian defined as

\[
H = -|0\rangle \langle 0| + \sum_{n=1}^{D-1} |n\rangle \langle n|.
\]  

Here the ground state is \( |0\rangle \) and has energy \( E_0 = -1 \), and the remaining states have an energy \( E_n = 1 \) for \( n \geq 1 \).

We again choose a full range of the domain of \( A_{k_0, k_1} \) such that \(-\pi/4 \leq x \leq \pi/4\), which means we choose \( \epsilon E_0 = -\pi/4 \) and \( \epsilon E_n = \pi/4 \) for \( n \geq 1 \), which means that \( \epsilon = \pi/4 \). In this case the measurement operators...
are according to (2) and (3) in the main text
\[ M_0 = |0⟩⟨0|, \]
\[ M_1 = \sum_{n=1}^{D-1} |n⟩⟨n|. \]
\[ (S30) \]

Consider again a uniformly distributed initial state
\[ |ψ_0⟩ = \frac{1}{\sqrt{D}} \sum_{n=0}^{D-1} |n⟩. \]
\[ (S31) \]

The probability of obtaining the ground state in this case is
\[ p_0 = \langle ψ_0|M_0^DM_0|ψ_0⟩ = \frac{1}{D}. \]
\[ (S32) \]

The measurement operators \[ (S30) \] are in this case projective operators and are orthogonal \[ M_0M_1 = 0. \] As such, after the first measurement outcome of either \[ M_0 \] or \[ M_1 \], all subsequent measurements are obtained with the same outcome. Therefore, the length of a failed sequence is length \[ T_{\text{fail}} = 1, \] corresponding to the \[ M_1 \] outcome alone.

Assuming that the corrective unitary \[ U_C \] produces a state with amplitude \( \sim 1/\sqrt{D} \) on the ground state after each failed sequence, the total number of measurements before obtaining an outcome \[ M_0 \] is
\[ T ≈ \frac{T_{\text{fail}}}{p_0} + 1 = D + 1, \]
\[ (S33) \]

where the additional 1 is the length of the successful sequence \[ M_0. \]

We note that in a similar way to \[ (S28) \], the golf course Hamiltonian can be a computationally hard problem by permuting the state labels such that the lowest energy state is not necessarily the state \[ |0⟩ \].

The above scaling was verified numerically by directly running the algorithm using the Hamiltonian \[ (S29) \], as shown in Fig. \[ (S11)c \]. We again see excellent agreement with the theoretical prediction of \[ (S33). \] Considering \[ D \] to be an exponentially large quantity, the time scaling of the problem is therefore exponential.

**Other problems**

Here we make a brief comment regarding more complex Hamiltonian problems that occur in combinatorial optimization problems such as 3SAT or MAXCUT. Such problems have a gap that is constant with respect to the system dimension, in a similar way to \[ (S28) \]. Such problems can be considered to be an intermediate case between the uniform density of states and the golf course energy landscape, since there may exist a high level of degeneracy particularly in the middle of the spectrum. We have seen that in both limiting cases, there is an exponential overhead. Hence we expect that in such combinatorial problems the time scaling remains exponential.

**CONVERGENCE ERROR OF THE ALGORITHM**

We now estimate the error attained by the algorithm during convergence of the algorithm. For a sufficiently large number of measurements, the algorithm converges to a \( k_0, k_1 \) such that \( x_{\text{max}}^k < ε_{\text{th}} \), so that the unitary \( U_{k_0k_1} = I \) as described in the previous section. The more the function \( A \) is peaked at an energy less than \( x_{\text{max}}^k = ε_{\text{th}} \), the better the fidelity is, since the Gaussian form tends to suppress high energy states. Hence the worst-case fidelity is when \( x_{\text{max}}^k = ε_{\text{th}} \), so that it barely satisfies the convergence threshold.

Now consider for simplicity a non-degenerate ground and first excited state, separated by an energy gap \( Δ \). Let us also parametrize \( ε_{\text{th}} = E_0 + δ \). According to Eq. (4) in the main text, the resulting unnormalized state is
\[ M_0^kM_1^k|ψ⟩ = \sum_{n} A_{k_0k_1}(ε_{E_0})(⟨E_n|ψ⟩|E_0⟩|E_n⟩) \]
\[ (S34) \]

where we can approximate the \( A \)-function by a Gaussian of form
\[ A_{k_0k_1}(x) ∝ e^{-K(x-x_{\text{max}}^k)^2}, \]
\[ (S35) \]

where \( K = k_0 + k_1 \) is the total number of measurements. The amplitude factors on the ground and first excited states are
\[ A_{k_0k_1}(ε_{E_0}) ∝ e^{-Kε^2δ^2} \]
\[ A_{k_0k_1}(ε_{E_1}) ∝ e^{-Kε^2(Δ-δ)^2}. \]
\[ (S36) \]

The fidelity with the ground state for the unnormalized state \[ (S34) \] is
\[ F = \frac{|A_{k_0k_1}(ε_{E_0})|⟨E_0|ψ⟩|^2}{\sum_{n=0}^{D-1} |A_{k_0k_1}(ε_{E_n})|⟨E_n|ψ⟩|^2} \]
\[ \approx \frac{1}{1 + \sum_{n=1}^{D} \frac{|A_{k_0k_1}(ε_{E_n})|⟨E_n|ψ⟩|^2}{|A_{k_0k_1}(ε_{E_0})|⟨E_0|ψ⟩|^2}}, \]
\[ (S37) \]

where \( D \) is the Hilbert space dimension. Due to the fact that \( A \) function is a Gaussian, assuming the initial coefficients \( ⟨E_n|ψ⟩ \) are of the same order, the fidelity can be estimated as
\[ F ≈ \frac{1}{1 + \frac{|A_{k_0k_1}(ε_{E_1})|⟨E_1|ψ⟩|^2}{|A_{k_0k_1}(ε_{E_0})|⟨E_0|ψ⟩|^2}} \]
\[ \approx 1 - \frac{A_{k_0k_1}(ε_{E_1})|⟨E_1|ψ⟩|^2}{A_{k_0k_1}(ε_{E_0})|⟨E_0|ψ⟩|^2}, \]
\[ (S38) \]

assuming that \( \frac{A_{k_0k_1}(ε_{E_1})|⟨E_1|ψ⟩|^2}{A_{k_0k_1}(ε_{E_0})|⟨E_0|ψ⟩|^2} ≪ 1 \). The error, or infidelity, is then
\[ ε = 1 - F \]
\[ \approx \frac{|⟨E_1|ψ⟩|^2}{|⟨E_0|ψ⟩|^2} e^{-Kε^2(Δ-2δ)} \]
\[ (S39) \]
In order to converge to the ground state, we must set $E_{\text{th}} < (E_0 + E_1)/2$, which using our variables corresponds to $\delta < \Delta/2$. When this is an equality, the exponential is equal to 1, and there is an equal suppression factor of both the ground and first excited states. Higher energy states are suppressed further.

**EXACT EVALUATION OF QUBIT IMAGINARY TIME EVOLUTION**

The qubit example (Example 1 in the main text) allows for another way of understanding the dynamics. In this case $M_0 M_1 \propto I$, so that a general sequence involving multiple applications of $U_C = X$ can be simplified as

$$\prod_{t=1}^{T} \left( U_{k_0}^{(t)} k_1 M_{m_t} \right) |\psi_0\rangle \propto M_0^k X^{N_C} |\psi_0\rangle \quad (S40)$$

where $k$ is a non-negative integer and $N_C$ is the number of the times $U_C = X$ is applied. It can be ensured that $k \geq 0$ because whenever $k_1 > k_0$ (according to the criterion $x_{k_k, k_1} > \epsilon E_{\text{th}} = 0$), $X$ is applied, $k_0 \leftrightarrow k_1$ are interchanged, since $XM_0 = M_1 X$, $X M_1 = M_0 X$. For example, we may simplify the sequence using these identities as

$$XM_1 M_1 XM_1 M_1 M_1 M_0 M_0 |\psi_0\rangle$$

$$= M_0 M_0 M_1 M_1 M_1 M_0 M_0 X^2 |\psi_0\rangle$$

$$\propto M_0 |\psi_0\rangle \quad (S41)$$

In this way, it is possible to always ensure that $k > 0$, which converges towards the ground state for $k \gg 1$. 
