LETTER • OPEN ACCESS

Boosting quantum amplitude exponentially in variational quantum algorithms

To cite this article: Thi Ha Kyaw *et al* 2024 *Quantum Sci. Technol.* **9** 01LT01

View the article online for updates and enhancements.

Quantum Science and Technology

LETTER

Boosting quantum amplitude exponentially in variational quantum algorithms

Thi Ha Kyaw[1,2,3,*] , Micheline B Soley[4,5,6,7,*] , Brandon Allen[6,7] , Paul Bergold[8] , Chong Sun[3] , Victor S Batista[6,7,9] and Alán Aspuru-Guzik[2,3,10,11]

1 LG Electronics Toronto AI Lab, Toronto, Ontario M5V 1M3, Canada
2 Department of Chemistry, University of Toronto, Toronto, Ontario M5G 1Z8, Canada
3 Department of Computer Science, University of Toronto, Toronto, Ontario M5S 2E4, Canada
4 Department of Chemistry, University of Wisconsin-Madison, 1101 University Ave., Madison, WI 53706, United States of America
5 Department of Physics, University of Wisconsin-Madison, 1150 University Ave., Madison, WI 53706, United States of America
6 Yale Quantum Institute, Yale University, PO Box 208334, New Haven, CT 06520-8263, United States of America
7 Department of Chemistry, Yale University, PO Box 208107, New Haven, CT 06520, United States of America
8 Department of Mathematics, University of Surrey, Guildford, United Kingdom
9 Energy Sciences Institute, Yale University, PO Box 27394, West Haven, CT 06516-7394, United States of America
10 Vector Institute for Artificial Intelligence, Toronto, Ontario M5S 1M1, Canada
11 Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada
* Authors to whom any correspondence should be addressed.

E-mail: thiha.kyaw@lge.com and micheline.soley@wisc.edu

Abstract

We introduce a family of variational quantum algorithms, which we coin as quantum iterative power algorithms (QIPAs), and demonstrate their capabilities as applied to global-optimization numerical experiments. Specifically, we demonstrate the QIPA based on a double exponential oracle as applied to ground state optimization of the H_2 molecule, search for the transmon qubit ground-state, and biprime factorization. Our results indicate that QIPA outperforms quantum imaginary time evolution (QITE) and requires a polynomial number of queries to reach convergence even with exponentially small overlap between an initial quantum state and the final desired quantum state, under some circumstances. We analytically show that there exists an exponential amplitude amplification at every step of the variational quantum algorithm, provided the initial wavefunction has non-vanishing probability with the desired state and that the unique maximum of the oracle is given by $\lambda_1 > 0$, while all other values are given by the same value $0 < \lambda_2 < \lambda_1$ (here λ can be taken as eigenvalues of the problem Hamiltonian). The generality of the global-optimization method presented here invites further application to other problems that currently have not been explored with QITE-based near-term quantum computing algorithms. Such approaches could facilitate identification of reaction pathways and transition states in chemical physics, as well as optimization in a broad range of machine learning applications. The method also provides a general framework for adaptation of a class of classical optimization algorithms to quantum computers to further broaden the range of algorithms amenable to implementation on current noisy intermediate-scale quantum computers.

1. Introduction

Quantum computers promise exponential speedup over classical counterparts in solving certain tasks [1]. When fault-tolerant general-purpose quantum computers become available, adiabatic state preparation and quantum phase estimation may become the standard quantum routines for determining the ground-state

energy of sophisticated physical Hamiltonians [2–5]. However, such schemes are very costly in terms of required overhead and hence are not suitable for the current era of noisy intermediate-scale quantum (NISQ) hardware [6–10]. This limitation of quantum computers today shifts central attention towards low-depth hybrid quantum–classical algorithms, known as NISQ algorithms [11–14]. The variational quantum eigensolver (VQE) [15, 16] serves as a prototypical example, as an algorithm that computes the expectation value of a Hamiltonian, which is measured on a quantum machine, resulting in a cost function with a set of variational parameters, which are optimized using classical computers. The process is repeated until the cost function reaches its local minimum.
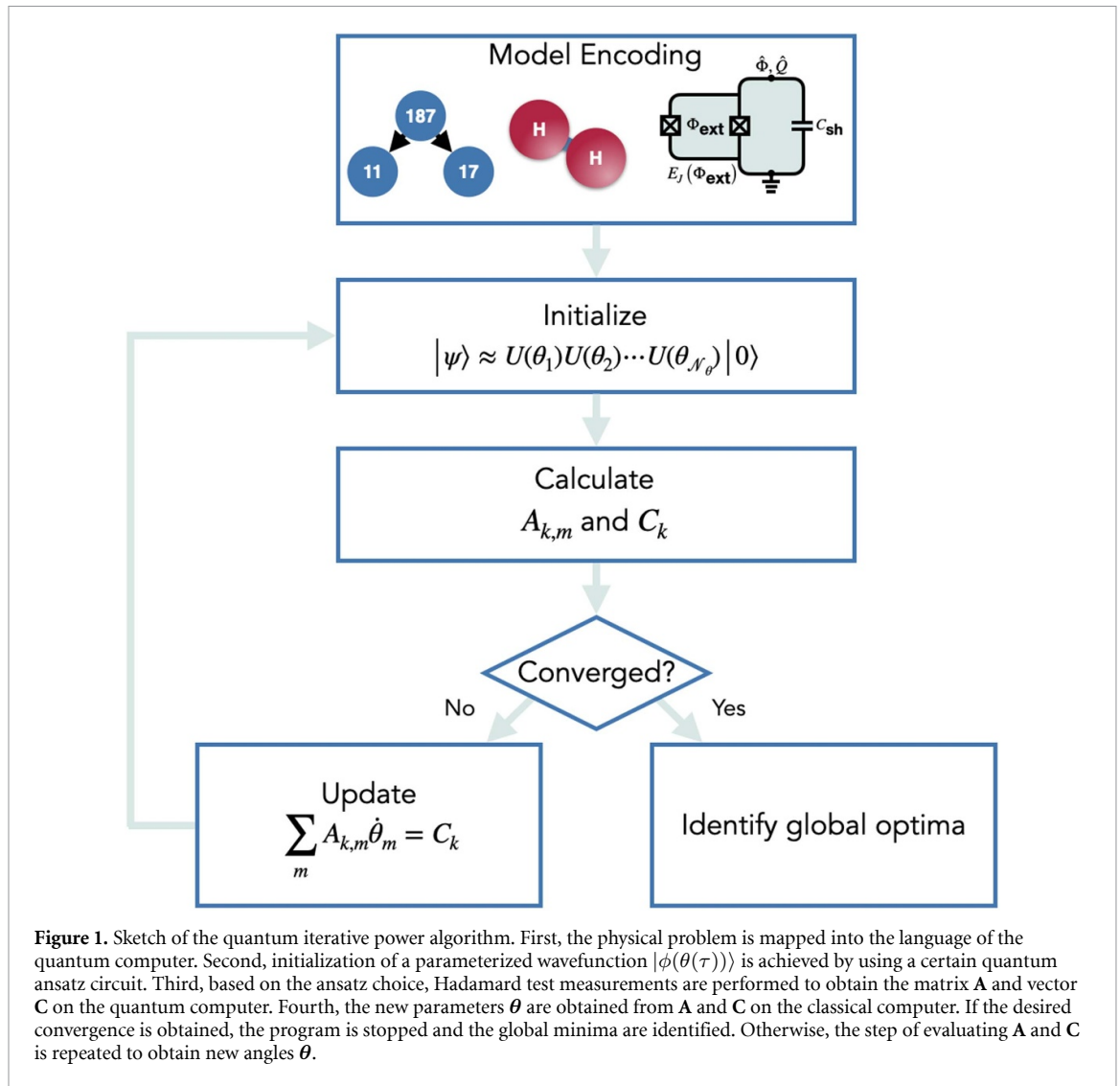
On the other hand, the variational quantum simulator [17] has been proposed for hybrid quantum–classical simulations of quantum dynamics based on the McLachlan's variational principle [18, 19], including quantum imaginary time evolution (QITE) to prepare ground states [18–20]. Here, we introduce the 'quantum iterative power algorithm' inspired by the variational quantum simulator to provide an accelerated method to the general problem of global optimization with near term quantum computers.

Global optimization is central to many important problems in science and engineering, from back-propagation in machine learning [21] and molecular geometry optimization/protein structure prediction [22, 23] to route planning and control of drone/unmanned aerial vehicles [24]. However, the brute force approach of considering each possible element of a search space often becomes computationally intractable. For example, identification of the optimal configuration of a protein faces Levinthal's paradox [25]—that the native configuration must be identified out of about 10^{300} possibilities. This has inspired a broad array of both classical [26] and quantum computing [27] optimizers. Recently, we have shown that tensor trains [28, 29] (also known as matrix product states) provide a way to vastly reduce the computational cost of exploring low-rank optimization cost functions, and have employed the approach to introduce an optimization algorithm that deterministically explores the full search space in data-compressed form, the tensor-train 'iterative power algorithm (IPA)' [30].

We recognize the strategy of tensor-train IPA can be implemented on quantum computers to enable global optimization of an even broader class of optimization problems. In tensor-train IPA, the optimization cost function of interest is taken to be a potential energy surface. A density is initialized in the potential energy surface, and an oracle is iteratively applied in a sifting approach akin to imaginary time propagation (with infinite mass) to localize the density as a delta function at the global minimum position. The expectation value of position then gives the location of the global minimum. Tensor-train IPA represents the density and potential energy surface as tensor trains to avoid calculation of the cost function everywhere in search space, which is efficient for representation of problems amenable to low-rank representations, such as prime factorization or molecular geometry optimization [30]. However, the tensor-train strategy faces the roadblock that highly-correlated systems cannot be efficiently represented in low-rank tensor-train format. In contrast, quantum computers excel in the simulation of highly-correlated systems, as the coupling or entanglement between qubits is limited only by the choice of ansatz [31].

The quantum iterative power algorithm (QIPA) takes advantage of the high degree of entanglement possible on quantum computers with a hybrid variational scheme. In standard variational approaches such as the VQE [15, 16], classical optimizers are used to determine the parameters of a quantum circuit, which are used to prepare trial wavefunctions measured to obtain expectation values. Analogously, the variational quantum simulator [17] evolves the parameters that define the time-evolved wavefunction by using a classical computer that integrates the Euler–Lagrange equation obtained from the Schrödinger equation with the McLachlan's variational principle. Parameters required by the Euler–Lagrange equation are obtained with a quantum circuit with a small number of quantum operations. QIPA generalizes the variational quantum approach to evolve an initial wavefunction such that the corresponding probability density (modulus squared of the wavefunction) becomes localized at the global minimum of a given cost function (see figure 1). As in IPA, the propagator of QIPA is not limited to the imaginary time quantum propagator enabling the use of other propagators that are maximal at the minimum of the cost function.

Figure 1 and algorithm 1 show the overall workflow of the QIPA. First, we select parameters $\theta_1, \theta_2, \ldots, \theta_{\mathcal{N}_\theta}$ corresponding to the initial state. Second, we use a quantum coprocessor to perform the Hadamard tests, resulting in $A_{k,m}$ and C_k of $\sum_m A_{k,m} \dot{\theta}_m = C_k$. There are a number of ways to get an approximate solution to the linear equation. Here, we used the conjugate gradient method without needing to invert the matrix $\mathbf{A}$ to get a solution $\dot{\theta}$, using a classical computer. Next, we update the parameters θ by the Euler method. Having updated the parameters, the process is iterated until convergence to obtain parameters θ corresponding to a distribution function localized at the global minimum.

Figure 1. Sketch of the quantum iterative power algorithm. First, the physical problem is mapped into the language of the quantum computer. Second, initialization of a parameterized wavefunction $|\phi(\theta(\tau))\rangle$ is achieved by using a certain quantum ansatz circuit. Third, based on the ansatz choice, Hadamard test measurements are performed to obtain the matrix **A** and vector **C** on the quantum computer. Fourth, the new parameters θ are obtained from **A** and **C** on the classical computer. If the desired convergence is obtained, the program is stopped and the global minima are identified. Otherwise, the step of evaluating **A** and **C** is repeated to obtain new angles θ.

Algorithm 1. Variational QIPA.

Require: Hamiltonian $\hat{H}$ and initial state $|\psi(0)\rangle = |0\rangle^{\otimes \mathcal{N}}$
  1: Start with an ansatz $|\phi(\boldsymbol{\theta}_0)\rangle = U(\boldsymbol{\theta}_0)|\bar{0}\rangle$ at time $\tau = 0$;
  2: Evaluate Hadamard tests to form the matrix $\mathbf{A}(\tau)$ and the vector $\mathbf{C}(\tau)$ (quantum computer subroutine);
  3: Compute an approximate solution ξ_τ of $\mathbf{A}(\tau)\dot{\boldsymbol{\theta}}(\tau) = \mathbf{C}(\tau)$ via the conjugate gradient method
    (classical computer subroutine);
  4: Update the parameter as $\boldsymbol{\theta}(\tau + \delta\tau) \leftarrow \boldsymbol{\theta}(\tau) + \xi_\tau \delta\tau$ and set $\tau \leftarrow \tau + \delta\tau$;
  5: Repeat steps 2–4 until $\tau = \tau_{\text{total}}$ or the convergence criteria is met;

---

2. Results

We are interested in a particular case of global optimization involving the search of the ground state of a Hamiltonian $\hat{H}$: a problem that is typically solved by the imaginary time propagation. QIPA can solve the same problem analogously by using $\hat{H}$ in the normalized oracle function $f(\hat{H}; \tau)$ that acts on the initial wavefunction $|\psi(0)\rangle$, as follows (onwards setting $\hbar = 1$):

$$
\begin{aligned}
|\psi(\tau)\rangle &= f(\hat{H}; \tau)|\psi(0)\rangle \\
&= \frac{U(\tau)|\psi(0)\rangle}{\sqrt{\langle U(\tau)\psi(0)|U(\tau)\psi(0)\rangle}},
\end{aligned}
\tag{1}
$$

where $U(\tau)$ is an arbitrary oracle function with maximum at the global minimum position of the potential energy surface M, or here the Hamiltonian $\hat{H}$. In the following, we show that oracles defined by concatenated exponential functions,

$$U(\tau) = U_n(\tau) = \beta_n(-\hat{H}\tau)) = e^{b_n\beta_{n-1}(-\hat{H}\tau)}, \tag{2}$$

with $n \geqslant 1$ the number of concatenated exponentials, $\beta_0(y) = y$, $\beta_1(y) = e^{b_1 y}$, $\beta_2(y) = e^{b_2 e^{b_1 y}}, \dots$ and real constants $b_1, \dots, b_n \neq 0$, provide effective QIPAs based on a generalization of the McLachlan's variational principle (SI appendix A.2.). For example, the oracle defined by the double-exponential $U_2(\tau) = e^{e^{-\tau \hat{H}}}$ is obtained by setting $n = 2$ and $b_2 = b_1 = 1$.

We remark that the choice of U_1 corresponds to the standard QITE, which is widely used in quantum Monte Carlo algorithms. References [18, 19] show that one can perform imaginary time evolution [20] with unitary gates defined by equation (1) with $n = 1$ that evolve the initial state according to the Wick-rotated Schrödinger equation: $\partial|\psi(\tau)\rangle/\partial\tau = -(\hat{H} - E_1(\tau))|\psi(\tau)\rangle$, where $E_1(\tau) = \langle\psi(\tau)|\hat{H}|\psi(\tau)\rangle$. Here, we introduce a family of near-term quantum algorithms defined by β_n with $n \geqslant 1$ that evolve the initial state according to the generalized Wick-like-rotated Schrödinger equation (SI appendix A.3.):

$$\frac{\partial}{\partial\tau}|\psi(\tau)\rangle = -\prod_{k=1}^{n} b_k \left(\hat{H}\exp\left(\hat{S}_{n-1}\right) - \mathrm{Re}\left\langle\hat{H}\exp\left(\hat{S}_{n-1}\right)\psi(\tau) \mid \psi(\tau)\right\rangle\right)|\psi(\tau)\rangle, \tag{3}$$

where $\hat{S}_{n-1} = \sum_{k=1}^{n-1} b_k\beta_{k-1}(-\hat{H}\tau)$.

With the choice $n = 2$ and $b_2 = b_1 = 1$ we arrive at a double-exponential function and the following Wick-like-rotated Schrödinger equation:

$$\frac{\partial}{\partial\tau}|\psi(\tau)\rangle = -\left(\hat{H}e^{-\hat{H}\tau} - E_2(\tau)\right)|\psi(\tau)\rangle, \tag{4}$$

with $E_2(\tau) = \langle\psi(\tau) \mid \hat{H}e^{-\hat{H}\tau} \mid \psi(\tau)\rangle$. According to the McLachlan's variational principle, when we constrain the equation of motion as such

$$\delta\left\|\left(\partial/\partial\tau + \left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau)\right]\right)|\psi(\tau)\rangle\right\|^2 = 0, \tag{5}$$
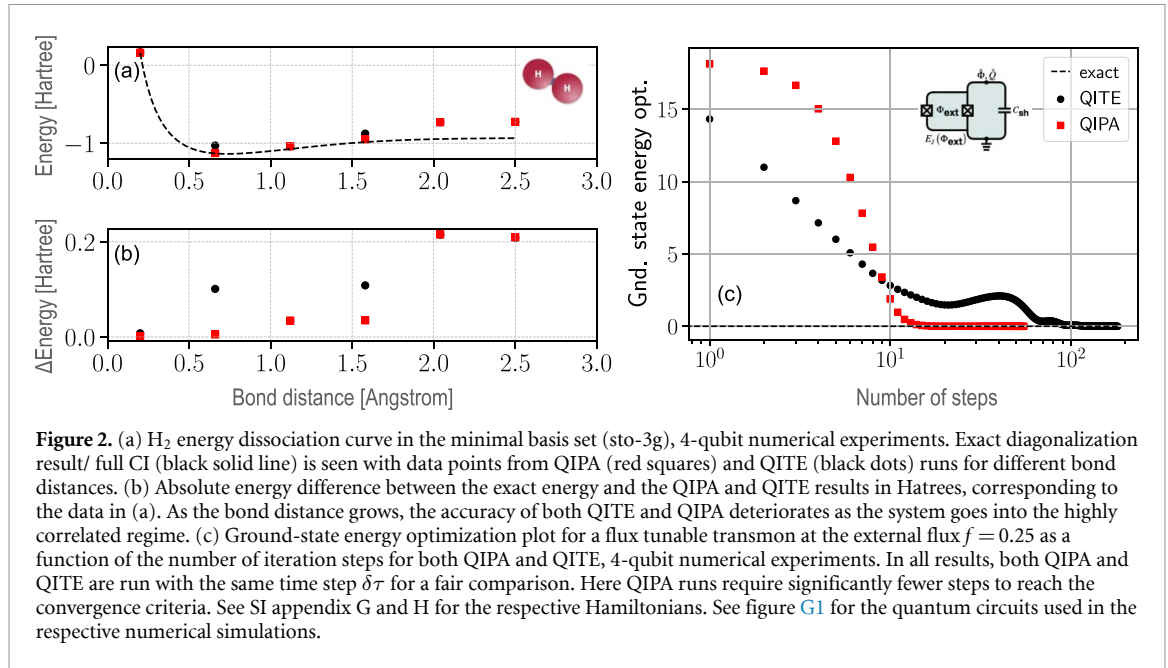
the result is equivalent in finding a solution of the linear equation: $\sum_m A_{k,m}\dot{\theta}_m = C_k$, where the entries of the symmetric and positive semi-definite matrix $\mathbf{A}$ and the right-hand side $\mathbf{C}$ can be computed on a quantum computer by deploying the Hadamard tests. The parameters $\boldsymbol{\theta}$ are updated with $\dot{\boldsymbol{\theta}}$ for a short time step $\delta\tau > 0$ according to the Euler method as $\boldsymbol{\theta}(\tau + \delta\tau) \approx \boldsymbol{\theta}(\tau) + \dot{\boldsymbol{\theta}}(\tau)\delta\tau$. The underlying assumption is that we can approximate $|\psi(\tau)\rangle$ by $|\phi(\boldsymbol{\theta}(\tau))\rangle = U(\theta_1(\tau))U(\theta_2(\tau))\cdots U(\theta_{\mathcal{N}_\theta}(\tau))|\bar{0}\rangle$, where $|\bar{0}\rangle = |0\rangle^{\otimes\mathcal{N}}$ and $U(\theta_1(\tau)),\dots,U(\theta_{\mathcal{N}_\theta}(\tau))$ are parameterized quantum circuits (PQCs), with $\boldsymbol{\theta} = (\theta_1,\dots,\theta_{\mathcal{N}_\theta})$ the corresponding real-valued parameter vector.

2.1. Exponential amplitude amplification

To precisely define what exponential amplitude amplification means, let us look at a general setting where we are interested to find a unique ground state $|\Psi\rangle$ of a quantum system $\hat{H}$, assuming no degeneracy. In variational quantum algorithms, one is interested to find the ground state as close as possible using hybrid quantum–classical approach and would end up obtaining an approximate state $|\Phi\rangle$, where $|\langle\Psi|\Phi\rangle| = \gamma$. It was recently shown that $\gamma \propto \exp(-\mathcal{N})$ [32] as the system size $\mathcal{N}$ grows for complex chemical molecular systems. In a specific condition that we are interested in (within the reach of the variational quantum ansatz), we prove analytically that γ can be amplified in exponentially less number of timesteps defined by the ratio $\iota = \lambda_{1,U_1}/\lambda_{2,U_1}$. The number of timesteps necessary to achieve more than 50% fidelity with the final desired quantum state is bounded by $k_{\mathrm{QIPA}}/k_{\mathrm{QITE}} \geqslant \frac{\log\iota}{\lambda_{2,U_1}(\iota-1)}$ (SI appendix D), which serves as a lower bound for all optimization problems in the manuscript, and that the proof can be generalized to the case of multiple global optima.

We are aware that local quantum Hamiltonian ground state problem is quantum Merlin Arthur (QMA)-complete and our approach does not change such problem's computational complexity class. We are merely pointing out that there is a special case with the proposed algorithm, where we are able to converge the solution in exponentially less number of steps as compared to the existing QITE program.

We remark that we used double-exponential oracle function as a particular working example. Other types of oracle functions such as $U(\tau) = \mathrm{sech}(\hat{H}\tau)$ can also be used (SI appendix F). The choice of an oracle function highly depends on the problem considered and the desired rate of convergence. The change in oracle function would result in the different convergence rate, with $k_{\mathrm{QIPA}}/k_{\mathrm{QITE}} \geqslant \varepsilon$, where $\varepsilon \ll 1$. Here, ε is an arbitrarily small number based on the choice of a new oracle.

Figure 2. (a) H$_2$ energy dissociation curve in the minimal basis set (sto-3g), 4-qubit numerical experiments. Exact diagonalization result/ full CI (black solid line) is seen with data points from QIPA (red squares) and QITE (black dots) runs for different bond distances. (b) Absolute energy difference between the exact energy and the QIPA and QITE results in Hatrees, corresponding to the data in (a). As the bond distance grows, the accuracy of both QITE and QIPA deteriorates as the system goes into the highly correlated regime. (c) Ground-state energy optimization plot for a flux tunable transmon at the external flux $f = 0.25$ as a function of the number of iteration steps for both QIPA and QITE, 4-qubit numerical experiments. In all results, both QIPA and QITE are run with the same time step $\delta\tau$ for a fair comparison. Here QIPA runs require significantly fewer steps to reach the convergence criteria. See SI appendix G and H for the respective Hamiltonians. See figure G1 for the quantum circuits used in the respective numerical simulations.

2.2. Resource estimate and error analysis

In general, for an $\mathcal{N}$-qubit system with Hamiltonian $\hat{H}$ with $\mathcal{N}_H \geqslant 1$ Pauli words and a parameterized wavefunction $|\phi(\boldsymbol{\theta})\rangle$ (where τ dependency $\boldsymbol{\theta}(\tau)$ is understood throughout) with $\mathcal{N}_\theta \geqslant 1$ parameters, the upper bound for the number of distinct measurements $\mathcal{N}_A$ required to obtain the matrix **A** for QIPA via the Hadamard test and the number of gates required are $\mathcal{N}_\theta(\mathcal{N}_\theta - 1)/2$ and $G_{\mathcal{N}_A} \geqslant \mathcal{N}_\theta$, respectively. Such an estimate can be understood as the number of times required to completely evaluate all the **A** matrix elements since **A** is symmetric. Moreover, to obtain the vector **C**, the number of measurements and gates required (assuming a second-order Taylor series expansion of the required function of the Hamiltonian) are $\mathcal{N}_\theta$ and $G_{\mathcal{N}_C} \geqslant \mathcal{N}_H + \mathcal{N}_H^2 + \mathcal{N}_H^3 + \mathcal{N}_\theta$, respectively. '>' sign in $G_{\mathcal{N}_A}$ and $G_{\mathcal{N}_C}$ holds when two-qubit gates are not parameterized, while '=' sign holds when they are parameterized. Assuming a polynomial scaling: $\mathcal{N}_H = \mathcal{O}(\mathcal{N}^h), \mathcal{N}_\theta = \mathcal{O}(\mathcal{N}^d)$, the leading order becomes $\mathcal{N}_A = \mathcal{O}(\mathcal{N}^d)$ and $\mathcal{N}_C = \mathcal{O}(\mathcal{N}^{\max(3h,d)})$, respectively. In comparison, in QITE, one needs $\mathcal{N}_A = \mathcal{O}(\mathcal{N}^d)$ and $\mathcal{N}_C = \mathcal{O}(\mathcal{N}^{\max(h,d)})$, with the same number of Hadamard test measurements required. In general, QIPA yields improved convergence in shorter times compared to QITE, requiring the same number of Hadamard test operations and a higher number of unitary gates. More importantly, we have also estimated that the error from the Taylor expansion causing the major difference between QITE and QIPA is given by $\epsilon = \sqrt{\Delta^2 \delta\tau^2 + \mathcal{O}(\delta\tau^3)} \leqslant \Delta\delta\tau + \mathcal{O}(\delta\tau^{3/2})$, where $\Delta^2 = \langle\Psi(t)|((1 + e^{-\hat{H}\delta\tau})^2/(\delta\tau^2) + 2(e^{-\hat{H}\delta\tau} - 1)\hat{H}/\delta\tau - (e^{-\hat{H}\delta\tau} - 2)\hat{H}^2)|\Psi(t)\rangle$ (SI appendix E).

2.3. Numerical experiments

Figure 2 shows that one can efficiently search for the ground-state energy of hydrogen molecule across various bond stretching distances with QIPA with fewer time steps than QITE. Results are shown in figure 2(a) comparing QIPA and QITE. For consistency and fair comparison, we use the same time step for each bond distance for both QIPA and QITE runs. The error difference between the exact energy obtained from full configuration interaction (CI) calculations and the QIPA and QITE results can be seen in figure 2(b). QIPA features less error for all bond distances considered even though the accuracy of both QITE and QIPA deteriorates as the system goes into the highly correlated regime. Alternative ansatz circuits could be considered for increased accuracy.

As shown in figure 2(c), both QITE and QIPA successfully minimize the energy of a given transmon circuit. Moreover, QIPA requires significantly fewer iterations than QITE for global optimization of the energy. These results constitute the successful implementation of two forms of imaginary-time-like evolution to perform quantum computer-aided design (QCAD), i.e. to think in a 'reverse-engineering way' to find out optimal quantum system parameters, in this case being classical circuit parameters of the superconducting circuit, given one desired energy spectrum of the quantum system. Previous methods [33–35] used VQE approach whereas ours is in the form of variational quantum simulator.

According to figures 3(a) and (b), we find QIPA identifies the prime factors of biprimes with fewer iterations than QITE for all integers factored. The number of required iterations varies depending on the biprime factored and the ansatz, with speedups shown here of up to 50%. In addition, QIPA successfully
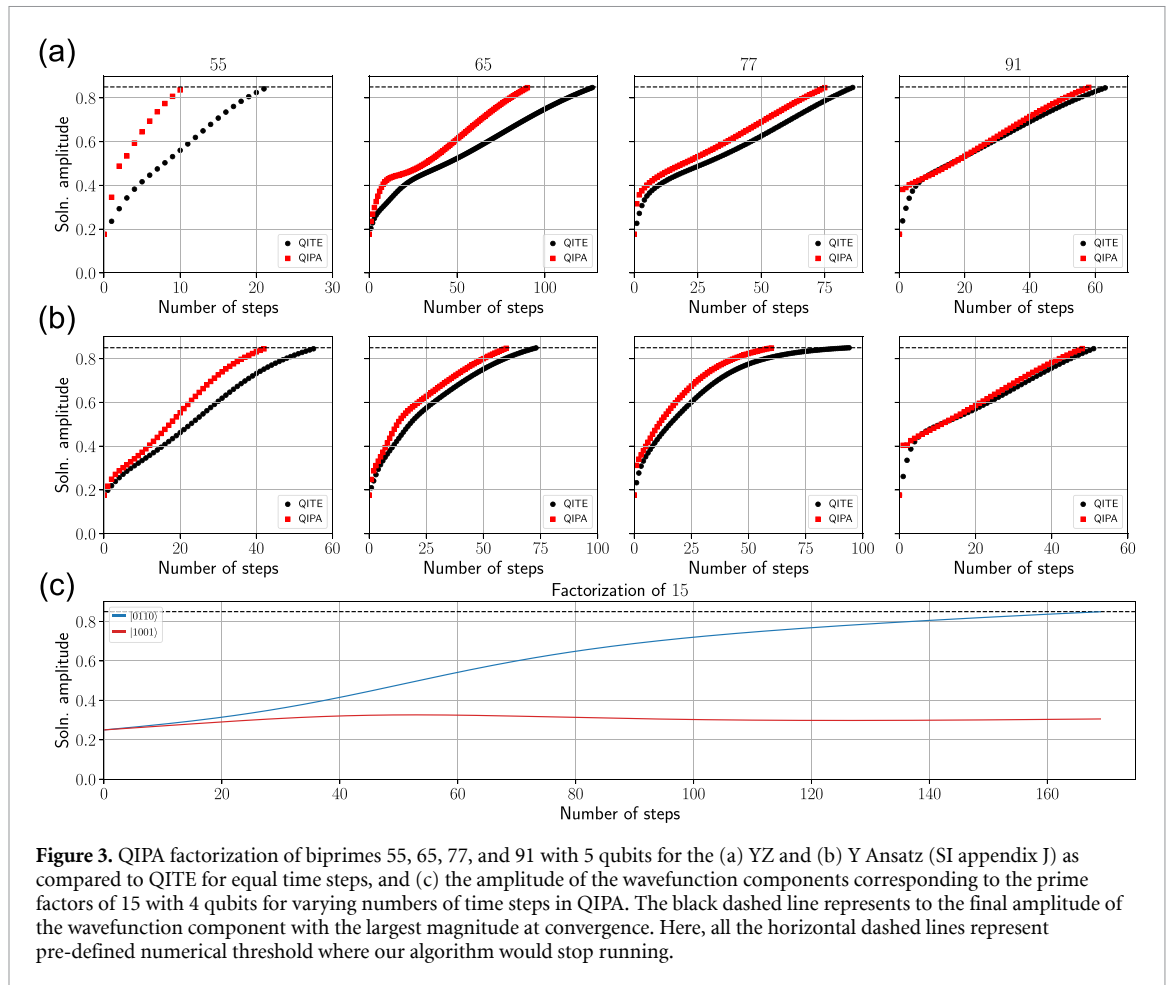
Figure 3. QIPA factorization of biprimes 55, 65, 77, and 91 with 5 qubits for the (a) YZ and (b) Y Ansatz (SI appendix J) as compared to QITE for equal time steps, and (c) the amplitude of the wavefunction components corresponding to the prime factors of 15 with 4 qubits for varying numbers of time steps in QIPA. The black dashed line represents to the final amplitude of the wavefunction component with the largest magnitude at convergence. Here, all the horizontal dashed lines represent pre-defined numerical threshold where our algorithm would stop running.

identifies both factors of biprimes. As expected, QIPA succeeds for stable integration with a small, converged time step, for which the twin solutions are readily identified as the two components of the final wavefunction with the largest and equal amplitude. Moreover, QIPA succeeds for *unstable* integration with a larger time step, for which the twin solutions are identified as the two largest components of the final wavefunction with *unequal* amplitude, as depicted in figure 3(c). To our knowledge this ability to identify multiple prime factors is unique among published quantum computing prime factorization results.

3. Discussion

In summary, we have presented a family of generalized imaginary-time-like near-term quantum algorithms which we coin the 'quantum iterative power algorithm,' inspired by its classical counterpart. (Plural 'algorithms' is used since depending on the choice of oracle function, the performance and behavior will differ. However, they all fall under the same family.) We have analyzed its convergence rate. One caveat is that since the proposed algorithm relies heavily on the ansatz circuit used, its convergence rate is difficult to discern in the generic case. We have also determined QIPA's estimated resource count as well as analytical error analysis, and demonstrated it can outperform the QITE while it reduces the number of required iterations, at the cost of a moderate increase in the number of gates. We note that even when the initial quantum state has an exponentially small overlap with the final target state, QIPA needs only a polynomial number of steps to reach convergence (SI appendix D). This is particularly important when starting with an initial state defined by a uniform superposition, or a low-rank reference state for a highly correlated system [32]. Furthermore, we have used the three numerical case studies—quantum chemical molecular simulation of the hydrogen molecule for various bond dissociation distances, quantum computer-aided design of a superconducting transmon, and finding optimal solutions for double prime factorization—to highlight how QIPA outperforms QITE. The simulation accuracy could be further improved by integrating with existing quantum error mitigation methods [17, 36–41].

We would like to point out that an additional consideration for such an algorithm besides the choice of the ansatz circuit is setting the right parameter for the time step $\delta\tau$ at each evolution step. A number of proposals [42, 43] suggest the use of an adaptive time step to overcome such an issue. However, there exists

an opportunity to develop a systematic way to adjust the time step $\delta\tau$ according to gradient descent, rather than with a heuristic argument on how to adjust $\delta\tau$. A major drawback with quantum imaginary-time-like evolution algorithm is that it involves constructing the matrix elements of **A** with NISQ hardware. Since we are working with noisy quantum hardware, any large fluctuation in the matrix elements would result in suboptimal angles $\boldsymbol{\theta}$. Future work will explore conditioning the matrix based on errors associated with performing the Hadamard tests.

Data availability statement

Acknowledgments

Appendix A. Quantum iterative power algorithm

This section outlines a quantum version of a generalized IPA method [30], the so-called quantum iterative power algorithm (QIPA) inspired by the imaginary time propagation (appendix A.1) and the variational quantum simulator [17]. We introduce a series of oracles defined by concatenated exponential functions which evolve the initial state according to a Wick-like-rotated Schrödinger equation obtained from a generalized McLachlan's variational principle. Then, we focus on the particular case of a double-exponential function that provides an efficient implementation of QIPA. For completeness, we include a section introducing the QITE method obtained from the McLachlan's variational principle.

A.1. Quantum imaginary time evolution

A.1.1. McLachlan variational principle approach
Let us consider a many-body Hermitian Hamiltonian $\hat{H}$.[12] Given an initial state $|\psi(0)\rangle$, non-unitary quantum imaginary time evolution (QITE) is defined by

$$|\psi(\tau)\rangle = \frac{e^{-\hat{H}\tau}|\psi(0)\rangle}{\sqrt{\langle\psi(0)|e^{-2\hat{H}\tau}|\psi(0)\rangle}}. \tag{A.1}$$

Note that the denominator is a normalization factor. Since a key to our proposal is the realization that there is nothing except the requirement that there be a continuous, integrable, strictly positive oracle $\hat{U}$ that is maximized at the global minima of $\hat{H}$ to prevent us from assuming a particular form of the oracle; we present QITE more generally in terms of non-unitary time evolution according to an oracle, as follows:

$$|\psi(\tau)\rangle = \frac{\beta\left(-\hat{H}\tau\right)|\psi(0)\rangle}{\sqrt{\langle\psi(0)|\beta\left(-2\hat{H}\tau\right)|\psi(0)\rangle}}, \tag{A.2}$$

where β could be any strictly increasing positive function. When $\beta(y) = e^y$, we recover imaginary time evolution, which corresponds directly to the Wick rotation ($\tau = -it$). With that choice of $\beta(y)$, we obtain that the above quantum state satisfies the Wick-like rotated Schrödinger equation:

---

[12] In practice, the Hamiltonian may be expressed as the weighted sum $H = \sum_i \lambda_i h_i$ with real coefficients λ_i and tensor products of Pauli matrices h_i, since Pauli matrices form a complete basis.

$$\frac{\partial |\psi(\tau)\rangle}{\partial \tau} = -\left(\hat{H} - E_1(\tau)\right)|\psi(\tau)\rangle, \tag{A.3}$$

where $E_1(\tau) = \langle\psi(\tau)|\hat{H}|\psi(\tau)\rangle$. Even though $|\psi(\tau)\rangle$ is a valid wavefunction that can be represented in a quantum computer, the non-unitary time evolution cannot be straightforwardly mapped to a quantum circuit based solely on unitary gates. Here, the McLachlan's variational principle [18, 19, 48, 49] comes to the rescue and demands that

$$\delta \left\| \left(\partial/\partial\tau + \left[\hat{H} - E_1(\tau)\right]\right)|\psi(\tau)\rangle \right\| = 0, \tag{A.4}$$

with $\|\cdot\|$ representing the L^2 norm and δ the functional derivative. In the following, we intend to simulate the action of non-unitary dynamics, (A.3), on a quantum computer via McLachlan's variational principle.

In variational quantum simulations, instead of directly encoding the quantum state $|\psi(\tau)\rangle$ at time τ, we approximate it with a PQC $|\psi(\tau)\rangle \approx |\phi(\boldsymbol{\theta}(\tau))\rangle$ with a real-valued parameter vector $\boldsymbol{\theta}(\tau) = (\theta_1(\tau), \theta_2(\tau), \ldots, \theta_{\mathcal{N}_\theta}(\tau))$. We assume that physically relevant quantum states span a restricted region of the full Hilbert space [50] for a given time interval, such that the trial state parameterized by $\boldsymbol{\theta}$ is sufficient to prepare a desired quantum state by applying a sequence of parameterized unitary gates $U(\boldsymbol{\theta}) = U_{\mathcal{N}_\theta}(\theta_{\mathcal{N}_\theta}) \cdots U_k(\theta_k) \cdots U_1(\theta_1)$ to the initial state $|\bar{0}\rangle = |0\cdots0\rangle$. Thus, we have $|\phi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|\bar{0}\rangle$, where $U(\boldsymbol{\theta})$ is referred to as the ansatz, and $U_k(\theta_k)$ is the kth unitary gate controlled by classical parameter θ_k. Here, we are only concerned with single- or two-qubit gates, which is sufficient for universal quantum computing.

According to McLachlan's variational principle, we require $\partial\|(\partial/\partial\tau + (\hat{H} - E_\tau))|\phi(\boldsymbol{\theta}(\tau))\rangle\|/\partial\dot{\theta}_k = 0$. We have

$$\begin{aligned}
&\delta\|(\partial/\partial\tau + [H - E_1(\tau)])|\phi(\boldsymbol{\theta}(\tau))\rangle\|^2 \\
&= \sum_{m,n} \frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_m} \frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_n} \dot{\theta}_m \dot{\theta}_n + \langle\phi(\boldsymbol{\theta}(\tau))|^2 \left(\hat{H} - E_1(\tau)\right)^2 |\phi(\boldsymbol{\theta}(\tau))\rangle \\
&\quad + \sum_m \frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_m} \left(\hat{H} - E_1(\tau)\right)|\phi(\boldsymbol{\theta}(\tau))\rangle\dot{\theta}_m \cdots \\
&\quad + \sum_m \langle\phi(\boldsymbol{\theta}(\tau))|\left(\hat{H} - E_1(\tau)\right)\frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_m}\dot{\theta}_m.
\end{aligned} \tag{A.5}$$

By differentiating with respect to $\dot{\theta}_k$, we obtain

$$\begin{aligned}
&\partial\|(\partial/\partial\tau + [\hat{H} - E_1(\tau)])|\phi(\boldsymbol{\theta}(\tau))\rangle\|^2/\partial\dot{\theta}_k \\
&= \sum_m \left(\frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_k} \frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_m} + \frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_m} \frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_k} \right)\dot{\theta}_m \\
&\quad + \frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_k}\hat{H}|\phi(\boldsymbol{\theta}(\tau))\rangle + \langle\phi(\boldsymbol{\theta}(\tau))|\hat{H}\frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_k},
\end{aligned} \tag{A.6}$$

where we use $\langle\phi(\boldsymbol{\theta}(\tau))\rangle\phi(\boldsymbol{\theta}(\tau)) = 1$. Finally, we conclude that

$$\partial\|\left(\partial/\partial\tau + [\hat{H} - E_1(\tau)]\right)|\phi(\boldsymbol{\theta}(\tau))\rangle\|/\partial\dot{\theta}_k = 0 \tag{A.7}$$

is equivalent to the following linear equation

$$\sum_m A_{k,m}\dot{\theta}_m = C_k \tag{A.8}$$

with

$$\begin{aligned}
A_{k,m} &= \text{Re}\left(\frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_k} \frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_m} \right) \quad \text{and} \\
C_k &= -\text{Re}\left(\frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_k}\hat{H}|\phi(\boldsymbol{\theta}(\tau))\rangle \right).
\end{aligned} \tag{A.9}$$

We note that

$$\frac{\partial \langle \phi(\boldsymbol{\theta}(\tau))|}{\partial \theta_k} \hat{H} | \phi(\boldsymbol{\theta}(\tau)) \rangle = \sum_{\beta} \lambda_{\beta} \frac{\partial \langle \phi(\boldsymbol{\theta}(\tau))|}{\partial \theta_k} h_{\beta} | \phi(\boldsymbol{\theta}(\tau)) \rangle, \qquad (A.10)$$

where the h_{β} are Pauli matrices and λ_{β} are corresponding coefficients. Hence, QITE reduces to solving the linear equation $\mathbf{A}\dot{\boldsymbol{\theta}} = \mathbf{C}$ for $\dot{\boldsymbol{\theta}}$ which could be accomplished by inversion of the matrix $\mathbf{A}$, as follows $\dot{\boldsymbol{\theta}} = \mathbf{A}^{-1}\mathbf{C}$. In our numerical simulations, however, we solve the linear equation via the conjugate gradient method [51] using a subroutine from the SciPy library with a tolerance for convergence of 10^{-6}. That approach by-passes the need of inverting the matrix $\mathbf{A}$.

A.1.2. Quantum circuit evaluations of A and C

We efficiently evaluate the components of $\mathbf{A}$ and $\mathbf{C}$ following [18, 52, 53] by implementing the Hadamard test with an additional ancilla qubit. Recall that $|\phi(\boldsymbol{\theta})\rangle = U_{\mathcal{N}_{\theta}}(\theta_{\mathcal{N}_{\theta}}) \cdots U_1(\theta_1)|\bar{0}\rangle$ and that in the variational ansatz circuit we are only concerned with single- and two-qubit unitary gates $U_n(\theta_n)$, namely rotational or controlled rotational gates. The required derivatives of the ansatz wavefunction are then determined as follows.

Suppose $U_n(\theta_n)$ is a single-qubit rotational gate $R_{\theta_n}^Z = e^{-i\theta_n \hat{Z}/2}$, with derivative $\partial U_n(\theta_n)/\partial \theta_n = -(i/2) \times \hat{Z} e^{-i\theta_n \hat{Z}/2}$. If $U_n(\theta_n)$ is a two-qubit controlled rotational gate $|0\rangle\langle 0| \otimes \hat{I} + |1\rangle\langle 1| \otimes R_{\theta_n}^Z$, the derivative is given by $\partial U_n(\theta_n)/\partial \theta_n = |1\rangle\langle 1| \otimes \partial R_{\theta_n}^Z/\partial \theta_n = (-i/2) \times |1\rangle\langle 1| \otimes \hat{Z} e^{-i\theta_n \hat{Z}/2}$. In our numerical simulations, we do not consider parameterized two-qubit gates for simplicity. Using the notation

$$\bar{U}_n = U_{\mathcal{N}_{\theta}}(\theta_{\mathcal{N}_{\theta}}) \cdots U_{n+1}(\theta_{n+1}) U_n(\theta_n) g_n U_{n-1}(\theta_{n-1}) \cdots U_2(\theta_2) U_1(\theta_1), \qquad (A.11)$$

we conclude that

$$\bar{U}_n = U_{\mathcal{N}_{\theta}}(\theta_{\mathcal{N}_{\theta}}) \cdots U_{n+1}(\theta_{n+1}) (2i \times \partial U_n(\theta_n)/\partial \theta_n) U_{n-1}(\theta_{n-1}) \cdots U_2(\theta_2) U_1(\theta_1). \qquad (A.12)$$

This implies

$$\frac{\partial |\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial \theta_n} = (-i/2) \bar{U}_n |\bar{0}\rangle. \qquad (A.13)$$

Consequently, we have

$$A_{k,m} = \text{Re}\left(\frac{\partial \langle \phi(\boldsymbol{\theta}(\tau))|}{\partial \theta_k} \frac{\partial |\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial \theta_m} \right) = \frac{1}{4}\text{Re}\left(\langle \bar{0}|\bar{U}_k^{\dagger}\bar{U}_m|\bar{0}\rangle \right) \qquad (A.14)$$

and

$$\begin{aligned}
C_k &= \text{Re}\left(-\frac{\partial \langle \phi(\boldsymbol{\theta}(\tau))|}{\partial \theta_k} \hat{H} | \phi(\boldsymbol{\theta}(\tau)) \rangle \right) \\
&= -\text{Re}\left(\sum_{\beta} \lambda_{\beta} \frac{\partial \langle \phi(\boldsymbol{\theta}(\tau))|}{\partial \theta_k} \hat{h}_{\beta} | \phi(\boldsymbol{\theta}(\tau)) \rangle \right) \\
&= -\frac{1}{2}\text{Re}\left(i \sum_{\beta} \lambda_{\beta} \langle \bar{0}|\bar{U}_k^{\dagger}\hat{h}_{\beta} U|\bar{0}\rangle \right).
\end{aligned} \qquad (A.15)$$

Since we are evaluating $\text{Re}(\langle \bar{0}|\bar{U}_k^{\dagger}\bar{U}_m|\bar{0}\rangle)$ and $\text{Re}(\langle \bar{0}|\bar{U}_k^{\dagger}\hat{h}_{\beta} U|\bar{0}\rangle)$, one can implement them on a quantum computer by carrying out the Hadamard tests shown in figure A1.

A.1.3. Parameter update

We make use of the Euler method (the first order Taylor series expansion) to update the variational parameters as

$$\boldsymbol{\theta}(\tau + \delta\tau) \simeq \boldsymbol{\theta}(\tau) + \dot{\boldsymbol{\theta}}(\tau)\delta\tau \approx \boldsymbol{\theta}(\tau) + \xi_{\tau}\delta\tau, \qquad (A.16)$$

where ξ_{τ} is the numerical solution to $\mathbf{A}(\tau)\dot{\boldsymbol{\theta}}(\tau) = \mathbf{C}(\tau)$. One needs to repeat this procedure $\mathcal{N}_T = \tau_{\text{total}}/\delta\tau$ times to simulate imaginary-time-like evolution. The difference between the above parameter update and the gradient descent method is that the latter uses

$$\boldsymbol{\theta}(\tau + \delta\tau) \simeq \boldsymbol{\theta}(\tau) + \mathbf{C}(\tau)\delta\tau, \qquad (A.17)$$

Figure A1. (a) Quantum circuit to evaluate $\mathrm{Re}(\langle\bar{0}|\bar{U}_i^\dagger\bar{U}_j|\bar{0}\rangle)$ as a probability of finding the ancillary qubit in 0. (b) Quantum circuit to evaluate $\mathrm{Re}(\langle\bar{0}|\bar{U}_i^\dagger\hat{h}U|\bar{0}\rangle)$ as a probability of finding the ancillary qubit in 0.
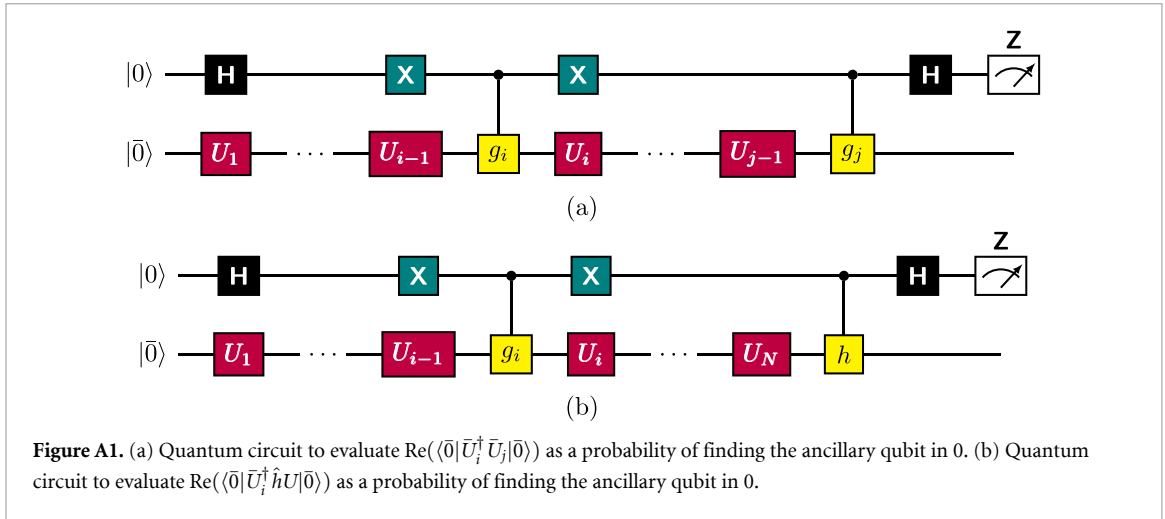
which only considers information about the average energy at each time step without taking into account of information about the ansatz circuit itself.

Appendix B. General formulation of a concatenated exponential cooling function

As mentioned in the main text, the main ingredient of QIPA is the choice of a suitable cooling function or oracle to quickly reach the optimal solution. Here, we show that oracles defined by the concatenated exponential functions introduced by equation (2) (main text) evolve the initial state according to the generalized Wick-like rotated Schrödinger equation introduced by equation (3) (main text). We apply the chain rule for derivatives to equation (2) (main text) to obtain

$$
\begin{aligned}
\frac{\partial}{\partial y}\beta_n(y) &= \beta_n(y)b_n\beta_{n-1}'(y) = \beta_n(y)b_n\beta_{n-1}(y)b_{n-1}\beta_{n-2}'(y)\\
&= \beta_n(y)b_n\beta_{n-1}(y)b_{n-1}\beta_{n-2}(y)\cdots b_1\beta_0'(y)\\
&= b_n\cdots b_1\exp(b_n\beta_{n-1}(y)+\cdots+b_1\beta_0(y))\\
&= \prod_{k=1}^n b_k\exp\left(\sum_{k=1}^n b_k\beta_{k-1}(y)\right).
\end{aligned}
\tag{B.1}
$$

For a given initial wave function $|\psi(0)\rangle$, let us introduce the auxiliary functions

$$
\begin{aligned}
g_{1,n}(\tau) &= U_n(\tau)|\psi(0)\rangle,\\
g_{2,n}(\tau) &= \sqrt{\langle U_n(\tau)\psi(0)\mid U_n(\tau)\psi(0)\rangle}.
\end{aligned}
\tag{B.2}
$$

Note that the normalized time-evolved wavefunction $|\psi(\tau)\rangle$ can be written as

$$
|\psi(\tau)\rangle = \frac{g_{1,n}(\tau)}{g_{2,n}(\tau)} = \frac{U_n(\tau)|\psi_0\rangle}{\sqrt{\langle U_n(\tau)\psi(0)\mid U_n(\tau)\psi(0)\rangle}}.
\tag{B.3}
$$

For the derivatives of the above functions we obtain

$$
\begin{aligned}
\frac{\partial}{\partial\tau}g_{1,n}(\tau) &= \frac{\partial}{\partial\tau}U_n(\tau)|\psi(0)\rangle\\
&= -\hat{H}U_n(\tau)b_n\beta_{n-1}'(-\hat{H}\tau)|\psi(0)\rangle = -\hat{H}b_n\beta_{n-1}'(-\hat{H}\tau)g_{1,n}(\tau),
\end{aligned}
$$

$$\frac{\partial}{\partial\tau}g_{2,n}(\tau) = \frac{1}{2}(\langle U_n(\tau)\psi(0) \mid U_n(\tau)\psi(0)\rangle)^{-1/2}\left(\left\langle \frac{\partial}{\partial\tau}U_n(\tau)\psi(0) \mid U_n(\tau)\psi(0)\right\rangle\right.$$

$$\left. + \left\langle U_n(\tau)\psi(0) \left| \frac{\partial}{\partial\tau}U_n(\tau)\psi(0)\right.\right\rangle\right)$$

$$= -\frac{1}{2}g_{2,n}(\tau)^{-1}a_n(\langle \hat{H}\beta'_{n-1}(-\hat{H}\tau)U_n(\tau)\psi(0) \mid U_n(\tau)\psi(0)\rangle$$

$$+ \langle U_n(\tau)\psi(0) \mid \hat{H}\beta'_{n-1}(-\hat{H}\tau) \mid U_n(\tau)\psi(0)\rangle),$$

$$\frac{\partial}{\partial\tau}|\psi(\tau)\rangle = \frac{\partial g_{1,n}(\tau)/\partial\tau}{g_{2,n}(\tau)} - \frac{g_{1,n}(\tau)\partial g_{2,n}(\tau)/\partial\tau}{g_{2,n}(\tau)^2}$$

$$= -\hat{H}b_n\beta'_{n-1}(-\hat{H}\tau)|\psi(\tau)\rangle + \frac{b_n}{2}(\langle \hat{H}\beta'_{n-1}(-\hat{H}\tau)\psi(\tau) \mid \psi(\tau)\rangle$$

$$+ \langle \psi(\tau) \mid \hat{H}\beta'_{n-1}(-\hat{H}\tau)\psi(\tau)\rangle)|\psi(\tau)\rangle$$

$$= -b_n(\hat{H}\beta'_{n-1}(-\hat{H}\tau) - \mathrm{Re}\langle \hat{H}\beta'_{n-1}(-\hat{H}\tau)\psi(\tau) \mid \psi(\tau)\rangle)|\psi(\tau)\rangle$$

$$= -\prod_{k=1}^{n}a_k\left(\hat{H}\exp\left(\sum_{k=1}^{n-1}b_k\beta_{k-1}(-\hat{H}\tau)\right)\right.$$

$$\left. -\mathrm{Re}\left\langle \hat{H}\exp\left(\sum_{k=1}^{n-1}b_k\beta_{k-1}(-\hat{H}\tau)\right)\psi(\tau) \mid \psi(\tau)\right\rangle\right)|\psi(\tau)\rangle, \tag{B.4}$$

which is identical to equation (3) (main text). As previously mentioned, the standard QITE is recovered for the choice $n = 1$, $b_1 = 1$ (standard exponential function, cf equation (A.3), as follows:

$$\frac{\partial}{\partial\tau}|\psi(\tau)\rangle = -(\hat{H} - E_1(\tau))|\psi(\tau)\rangle, \quad E_1(\tau) = \langle \psi(\tau) \mid \hat{H} \mid \psi(\tau)\rangle. \tag{B.5}$$

Moreover, for the choice $n = 2$, $b_2 = 1$, $b_1 = 1$ (double-exponential function) we obtain

$$\frac{\partial}{\partial\tau}|\psi(\tau)\rangle = -(\hat{H}e^{-\hat{H}\tau} - E_2(\tau))|\psi(\tau)\rangle,$$

$$E_2(\tau) = \langle \psi(\tau) \mid \hat{H}e^{-\hat{H}\tau} \mid \psi(\tau)\rangle.$$

We see that the resulting equation for double-exponential QIPA takes a similar form to the original Wick-rotated equation of motion, but with more rapid convergence to the ground state of $\hat{H}$.
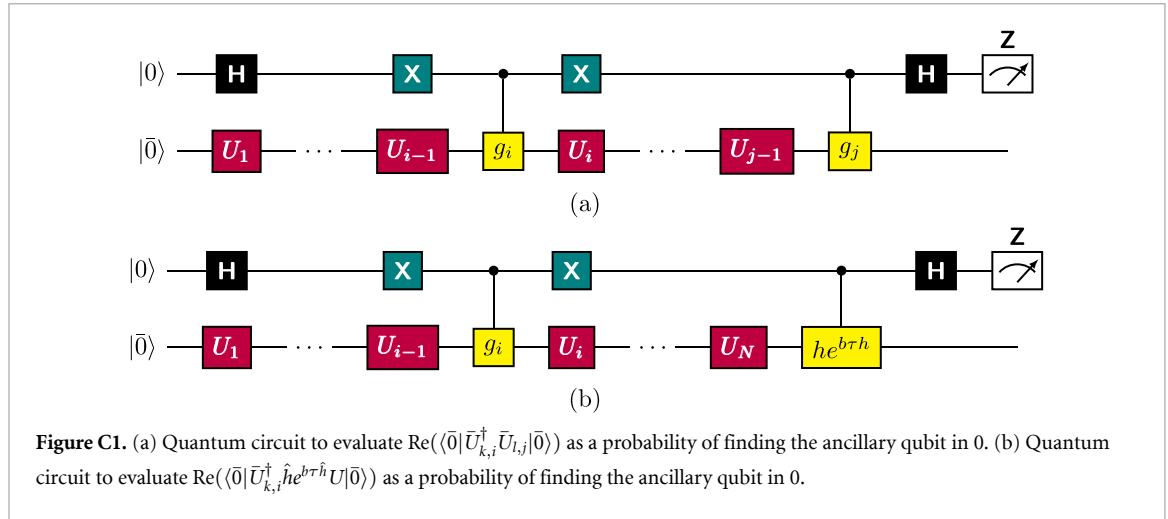
Appendix C. McLachlan's variational principle for the double-exponential oracle function

The McLachlan's variational principle is equivalent to the following minimization problem:

$$\delta\left\|\left(\partial/\partial\tau + \left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau)\right]\right)|\psi(\tau)\rangle\right\|^2 = 0. \tag{C.1}$$

Since $\boldsymbol{\theta}$ is real, we have

$$\left\|\left(\partial/\partial\tau + \left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau)\right]\right)|\phi(\boldsymbol{\theta}(\tau))\rangle\right\|^2$$

$$= \sum_{m,n}\frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_m}\frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_n}\dot{\theta}_m\dot{\theta}_n + \langle\phi(\boldsymbol{\theta}(\tau))|^2\left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau)\right]^2|\phi(\boldsymbol{\theta}(\tau))\rangle\dots$$

$$+ \sum_m\frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_m}\left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau)\right]|\phi(\boldsymbol{\theta}(\tau))\rangle\dot{\theta}_m$$

$$+ \sum_m\langle\phi(\boldsymbol{\theta}(\tau))|\left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau)\right]\frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_m}\dot{\theta}_m. \tag{C.2}$$

Figure C1. (a) Quantum circuit to evaluate $\mathrm{Re}(\langle\bar{0}|\bar{U}_{k,i}^{\dagger}\bar{U}_{l,j}|\bar{0}\rangle)$ as a probability of finding the ancillary qubit in 0. (b) Quantum circuit to evaluate $\mathrm{Re}(\langle\bar{0}|\bar{U}_{k,i}^{\dagger}\hat{h}e^{b\tau\hat{h}}U|\bar{0}\rangle)$ as a probability of finding the ancillary qubit in 0.

Differentiating equation (C.2) with respect to $\dot{\theta}_k$, we obtain

$$
\begin{aligned}
\partial \left\| \left(\partial/\partial\tau + \left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau) \right] \right) |\phi(\boldsymbol{\theta}(\tau))\rangle \right\|^2 / \partial\dot{\theta}_k \\
= \sum_m \left(\frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_k} \frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_m} + \frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_m} \frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_k} \right) \dot{\theta}_m \cdots \\
+ \frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_k} \left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau) \right] |\phi(\boldsymbol{\theta}(\tau))\rangle \\
+ \langle\phi(\boldsymbol{\theta}(\tau))| \left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau) \right] \frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_k}.
\end{aligned}
\tag{C.3}
$$

Consequently, since $\langle\phi(\boldsymbol{\theta}(\tau))\,|\,\phi(\boldsymbol{\theta}(\tau))\rangle = 1$, we conclude that

$$
\left\| \left(\partial/\partial\tau + \left[\hat{H}e^{-\hat{H}\tau} - E_2(\tau) \right] \right) |\phi(\boldsymbol{\theta}(\tau))\rangle \right\| = 0
\tag{C.4}
$$

is equivalent to the following linear equation

$$
\sum_m A_{k,m}\dot{\theta}_m = C_k
\tag{C.5}
$$

with

$$
\begin{aligned}
A_{k,m} &= \mathrm{Re}\left(\frac{\partial\langle\phi(\boldsymbol{\theta}(\tau)|}{\partial\theta_k} \frac{\partial|\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_m} \right) \quad \text{and,} \\
C_k &= -\mathrm{Re}\left(\frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_k} \hat{H}e^{-\hat{H}\tau}|\phi(\boldsymbol{\theta}(\tau))\rangle \right).
\end{aligned}
\tag{C.6}
$$

Therefore, as in the case of the standard exponential oracle, the McLachlan's principle reduces to a linear system of equations.

C.1. Quantum circuit evaluations of A and C

The matrix **A** and the vector **C** can be obtained from the circuits shown in figure C1. As seen from the above equation (C.6), the main difference between QIPA and QITE, in terms of comparing the use of double-exponential and exponential cooling functions, is the presence of $e^{-\hat{H}\tau}$ in C_k. Here, we approximate the exponential by its Taylor series expansion to the second order, as follows:

$$
\begin{aligned}
C_k &= \mathrm{Re}\left(\langle\bar{0}|\bar{U}_k^{\dagger}\hat{H}U|\bar{0}\rangle \right) - \tau\mathrm{Re}\left(\langle\bar{0}|\bar{U}_k^{\dagger}\hat{H}^2U|\bar{0}\rangle \right) \\
&\quad + \frac{\tau^2}{2}\mathrm{Re}\left(\langle\bar{0}|\bar{U}_k^{\dagger}\hat{H}^3U|\bar{0}\rangle \right) + \mathcal{O}\left(\hat{H}^4 \right).
\end{aligned}
\tag{C.7}
$$

Appendix D. Convergence rate analysis

In [30] we presented a formal proof of convergence for IPA applied to discrete optimization problems with a single global minimum and showed that the number of IPA steps necessary to amplify the amplitude of the global minimum to a value higher than $1/2 = 50\%$ scales logarithmically with the number $n \geqslant 2$ of possible states. More precisely, assuming that the initial wavefunction is given by the uniform superposition over all elements of the finite search space and that the unique maximum of the oracle is given by $\lambda_1 > 0$, while all other values are given by the same value $0 < \lambda_2 < \lambda_1$ (here λ can be taken as eigenvalues of the problem Hamiltonian), it was shown that after $k \geqslant 1$ IPA iteration steps, the ratio between the minimum and the maximum amplitude of the evolved density function ρ_k (modulus squared of the wavefunction) is given by

$$\frac{\rho_{k,\min}}{\rho_{k,\max}} = \left(\frac{\lambda_2}{\lambda_1}\right)^k. \tag{D.1}$$

In the following we compare the ratio $\mu = \lambda_2/\lambda_1$ for the exponential oracle $U_1(\tau) = e^{-\tau\hat{H}}$ and the double-exponential oracle $U_2(\tau) = e^{e^{-\tau\hat{H}}}$.

Let us assume that $\hat{H} = M$ (quantum propagator with infinite mass) and that the potential energy surface M attains its unique global minimum at position x^*. The global maximum of the oracle function $U_1(\tau)$ is then given by

$$\lambda_{1,U_1} = e^{-\tau M(x^*)}, \tag{D.2}$$

where the index U_1 is used to indicate that we consider the exponential oracle. Since λ_{1,U_1} is the unique global maximum and all other values of the oracle are equal, it also follows that

$$\lambda_{2,U_1} = e^{-\tau M(x_1)}, \tag{D.3}$$

where $x_1 \neq x^*$ is one of the remaining points of the search space. Moreover, for the double-exponential oracle $U_2(\tau)$ we obtain

$$\lambda_{1,U_2} = e^{e^{-\tau M(x^*)}} = e^{\lambda_{1,U_1}} \quad \text{and} \quad \lambda_{2,U_2} = e^{e^{-\tau M(x_1)}} = e^{\lambda_{2,U_1}}, \tag{D.4}$$

where the index U_2 is used to indicate that we are now considering the double-exponential. Hence we get

$$\mu_{\mathrm{QIPA}} = \frac{e^{\lambda_{2,U_1}}}{e^{\lambda_{1,U_1}}} = e^{\lambda_{2,U_1} - \lambda_{1,U_1}} = e^{-\left(\lambda_{1,U_1} - \lambda_{2,U_1}\right)}, \tag{D.5}$$

which shows that the ratio of the double-exponential oracle decays exponentially in the gap $\lambda_{1,U_1} - \lambda_{2,U_1}$ of the exponential oracle. The analysis presented in [30] therefore yields that the minimum number of required QIPA iterations (with double-exponential oracle) is given by

$$k_{\mathrm{QIPA}} \geqslant -\frac{\log(n-1)}{\log(\mu_{\mathrm{QIPA}})} = \frac{\log(n-1)}{\lambda_{1,U_1} - \lambda_{2,U_1}}, \tag{D.6}$$

which scales logarithmically with the size of the search space n and inversely with the difference $\lambda_{1,U_1} - \lambda_{2,U_1}$. Recall that this must be compared with the minimum number of required IPA iterations (exponential oracle), given by

$$k_{\mathrm{QITE}} \geqslant \frac{\log(n-1)}{\log(\lambda_{1,U_1}/\lambda_{2,U_1})}, \tag{D.7}$$

which is significantly larger if $\lambda_{1,U_1}/\lambda_{2,U_1} \approx 1$ and $\lambda_{1,U_1} - \lambda_{2,U_1} \gg 1$.

Appendix E. Analytical error analysis

Concerning the error of our proposed QIPA as compared to QITE, the difference lies in equation (C.6). To be precise,

$$C_{k,\mathrm{QIPA}} = -\mathrm{Re}\left(\frac{\partial\langle\phi(\boldsymbol{\theta}(\tau))|}{\partial\theta_k}\hat{H}e^{-\hat{H}\tau}|\phi(\boldsymbol{\theta}(\tau))\rangle\right), \tag{E.1}$$

while

$$C_{k,\text{QITE}} = -\text{Re}\left(\frac{\partial \langle \phi(\boldsymbol{\theta}(\tau))|}{\partial \theta_k} \hat{H}|\phi(\boldsymbol{\theta}(\tau))\rangle\right). \tag{E.2}$$

In that sense, the main difference lies in $e^{-\hat{H}\tau}|\phi(\boldsymbol{\theta}(\tau))\rangle$ and $|\phi(\boldsymbol{\theta}(\tau))\rangle$. Since we are taking the Taylor expansion of the exponential term in $C_{k,\text{QIPA}}$, the error can be calculated by the Euclidean distance between the exact imaginary time-evolved state and the approximate one as

$$\epsilon = ||e^{-\hat{H}\delta\tau}|\Psi(t)\rangle - \left(1 - \delta\tau\hat{H} + (\delta\tau)^2\hat{H}^2/2 + \mathcal{O}(\delta\tau^3)\right)|\Psi(t)\rangle||, \tag{E.3}$$

with $|||\psi\rangle|| = \sqrt{\langle\psi\rangle\psi}$ being the L^2 norm. By expanding the norm and using the triangle inequality, we arrive at

$$\epsilon = \sqrt{\Delta^2\delta\tau^2 + \mathcal{O}(\delta\tau^3)} \leqslant \Delta\delta\tau + \mathcal{O}\left(\delta\tau^{3/2}\right), \tag{E.4}$$

where

$$\Delta^2 = \langle\Psi(t)|\left(\left(1 + e^{-\hat{H}\delta\tau}\right)^2/\left(\delta\tau^2\right) + 2\left(e^{-\hat{H}\delta\tau} - 1\right)\hat{H}/\delta\tau - \left(e^{-\hat{H}\delta\tau} - 2\right)\hat{H}^2\right)|\Psi(t)\rangle. \tag{E.5}$$

Appendix F. McLachlan's variational principle for hyperbolic secant oracle function

The time-evolved wavefunction can be expressed as

$$|\psi(\tau)\rangle = \frac{g_{1,n}(\tau)}{g_{2,n}(\tau)}, \tag{F.1}$$

where

$$g_{1,n}(\tau) = U_n(\tau)|\psi(0)\rangle, \tag{F.2}$$

$$g_{2,n}(\tau) = \sqrt{\langle U_n(\tau)\psi(0)|U_n(\tau)\psi(0)\rangle}. \tag{F.3}$$

The evolution of the wavefunction is then

$$\frac{\partial|\psi(\tau)\rangle}{\partial\tau} = \frac{\partial g_{1,n}(\tau)/\partial\tau}{g_{2,n}(\tau)} - \frac{g_{1,n}(\tau)\partial g_{2,n}(\tau)/\partial\tau}{g_{2,n}(\tau)}, \tag{F.4}$$

where, without specifying the oracle $U_n(\tau)$, the required partial derivatives are

$$\frac{\partial g_{1,n}(\tau)}{\partial\tau} = \frac{\partial}{\partial\tau}U_n(\tau)\psi(0) \tag{F.5}$$

$$\begin{aligned}\frac{\partial g_{2,n}(\tau)}{\partial\tau} &= \frac{1}{2}\left(\langle U_n(\tau)\psi(0)|U_n(\tau)\psi(0)\rangle\right)^{-1/2}\left(\left\langle\frac{\partial}{\partial\tau}U_n(\tau)\psi(0)|U_n(\tau)\psi(0)\right\rangle\right.\\ &\quad\left.+ \left\langle U_n(\tau)\psi(0)|\frac{\partial}{\partial\tau}U_n(\tau)\psi(0)\right\rangle\right)\\ &= \frac{1}{2}g_{2,n}^{-1}\left(\left\langle\frac{\partial g_{1,n}(\tau)}{\partial\tau}|g_{1,n}(\tau)\right\rangle + \left\langle g_{1,n}(\tau)|\frac{\partial g_{1,n}(\tau)}{\partial\tau}\right\rangle\right)\\ &= \frac{1}{2}g_{2,n}^{-1}\left(\left\langle\frac{\partial g_{1,n}(\tau)}{\partial\tau}|g_{1,n}(\tau)\right\rangle + \left\langle g_{1,n}(\tau)|\frac{\partial g_{1,n}(\tau)}{\partial\tau}\right\rangle\right).\end{aligned} \tag{F.6}$$

In the special case that

$$U_n = \text{sech}\left(\hat{H}\tau\right) = \frac{2}{e^{\hat{H}\tau} + e^{-\hat{H}\tau}}, \tag{F.7}$$

the required functions are

$$g_{1,n}(\tau) = \frac{2}{e^{\hat{H}\tau} + e^{-\hat{H}\tau}}|\psi(0)\rangle \tag{F.8}$$

$$g_{2,n}(\tau) = \sqrt{\left\langle \frac{2}{e^{\hat{H}\tau} + e^{-\hat{H}\tau}} \psi(0) \Big| \frac{2}{e^{\hat{H}\tau} + e^{-\hat{H}\tau}} \psi(0) \right\rangle}$$

$$= 2\sqrt{\left\langle \frac{1}{e^{\hat{H}\tau} + e^{-\hat{H}\tau}} \psi(0) \Big| \frac{1}{e^{\hat{H}\tau} + e^{-\hat{H}\tau}} \psi(0) \right\rangle} \tag{F.9}$$

such that the gradient of the first function is

$$\frac{\partial g_{1,n}(\tau)}{\partial \tau} = \frac{\partial}{\partial \tau} \operatorname{sech}(\hat{H}\tau) \psi(0)$$

$$= \frac{2\hat{H}\left(e^{\hat{H}\tau} - e^{-\hat{H}\tau}\right)}{\left(e^{-\hat{H}\tau} + e^{\hat{H}\tau}\right)^2} \psi(0)$$

$$= \frac{\hat{H}\left(e^{\hat{H}\tau} - e^{-\hat{H}\tau}\right)}{e^{-\hat{H}\tau} + e^{\hat{H}\tau}} \operatorname{sech}(\hat{H}\tau) \psi(0)$$

$$= \frac{\hat{H}\left(e^{\hat{H}\tau} - e^{-\hat{H}\tau}\right)}{e^{-\hat{H}\tau} + e^{\hat{H}\tau}} g_{1,n}(\tau)$$

$$= H \tanh(\hat{H}\tau) g_{1,n}(\tau) \tag{F.10}$$

and the gradient of the second function is

$$\frac{\partial g_{2,n}(\tau)}{\partial \tau} = \frac{1}{2} g_{2,n}^{-1} \left(\left\langle \hat{H} \tanh(\hat{H}\tau) g_{1,n}(\tau) | g_{1,n}(\tau) \right\rangle + \left\langle g_{1,n}(\tau) | \hat{H} \tanh(\hat{H}\tau) g_{1,n}(\tau) \right\rangle \right)$$

$$= g_{2,n}^{-1} \operatorname{Re}\left(\left\langle g_{1,n}(\tau) \big| \hat{H} \tanh(\hat{H}\tau) \big| g_{1,n}(\tau) \right\rangle \right). \tag{F.11}$$

The evolution of the wavefunction is then

$$\frac{\partial |\psi(\tau)\rangle}{\partial \tau} = \frac{\hat{H} \tanh(\hat{H}\tau) g_{1,n}(\tau)}{g_{2,n}(\tau)} - \frac{g_{1,n}(\tau) g_{2,n}^{-1} \operatorname{Re}\left\langle g_{1,n}(\tau) \big| \hat{H} \tanh(\hat{H}\tau) \big| g_{1,n}(\tau) \right\rangle}{g_{2,n}(\tau)}$$

$$= \hat{H} \tanh(\hat{H}\tau) |\psi(\tau)\rangle - \operatorname{Re}\left\langle \left\langle \psi(\tau) \big| \hat{H} \tanh(\hat{H}\tau) \big| \psi(\tau) \right\rangle \right\rangle |\psi(\tau)\rangle$$

$$= \hat{H} \tanh(\hat{H}\tau) |\psi(\tau)\rangle - \operatorname{Re}\left[\left\langle \hat{H} \tanh(\hat{H}\tau) \right\rangle \right] |\psi(\tau)\rangle$$

$$= \hat{H} \tanh(\hat{H}\tau) |\psi(\tau)\rangle - \left\langle \hat{H} \tanh(\hat{H}\tau) \right\rangle |\psi(\tau)\rangle$$

$$= \left(\hat{H} \tanh(\hat{H}\tau) - E_3(\tau) \right) |\psi(\tau)\rangle, \quad \text{and}$$

$$E_3(\tau) = \left\langle \psi(\tau) \big| \hat{H} \tanh(\hat{H}\tau) \big| \psi(\tau) \right\rangle, \tag{F.12}$$

in direct analogy to the exponential and double-exponential oracles. The McLachlan variational principle corresponding to the resulting minimization problem is

$$\delta \left\| \left(\partial/\partial\tau - \left[\hat{H} \tanh(\hat{H}\tau) - E_3 \right] \right) |\psi(\tau)\rangle \right\|^2 = 0, \tag{F.13}$$

which, again in direct analogy to the exponential and double-exponential oracles, is equivalent to the linear equation

$$\sum_m A_{k,m} \dot{\theta}_m = C_k, \tag{F.14}$$

where

$$A_{k,m} = \operatorname{Re}\left(\frac{\partial \langle \phi(\boldsymbol{\theta}(\tau))|}{\partial \theta_k} \frac{\partial |\phi(\boldsymbol{\theta}(\tau))\rangle}{\partial \theta_m} \right), \tag{F.15}$$

$$C_k = \operatorname{Re}\left(\frac{\partial \langle \phi(\boldsymbol{\theta}(\tau))|}{\partial \theta_k} H \tanh(H\tau) |\phi(\boldsymbol{\theta}(\tau))\rangle \right). \tag{F.16}$$

Appendix G. Ansatz circuits and example quantum circuit to evaluate a matrix element A

Figure G1. Ansatz circuit used for (a) the H_2 dissociation curve (b) Y ansatz (c) YZ ansatz used in factorization of number 15 and the transmon qubit ground-state search, respectively. For (b) and (c), the example circuits are shown. The actual number of qubits varies in regards to how big a bi-prime is and how many transmons are present in a simulation.
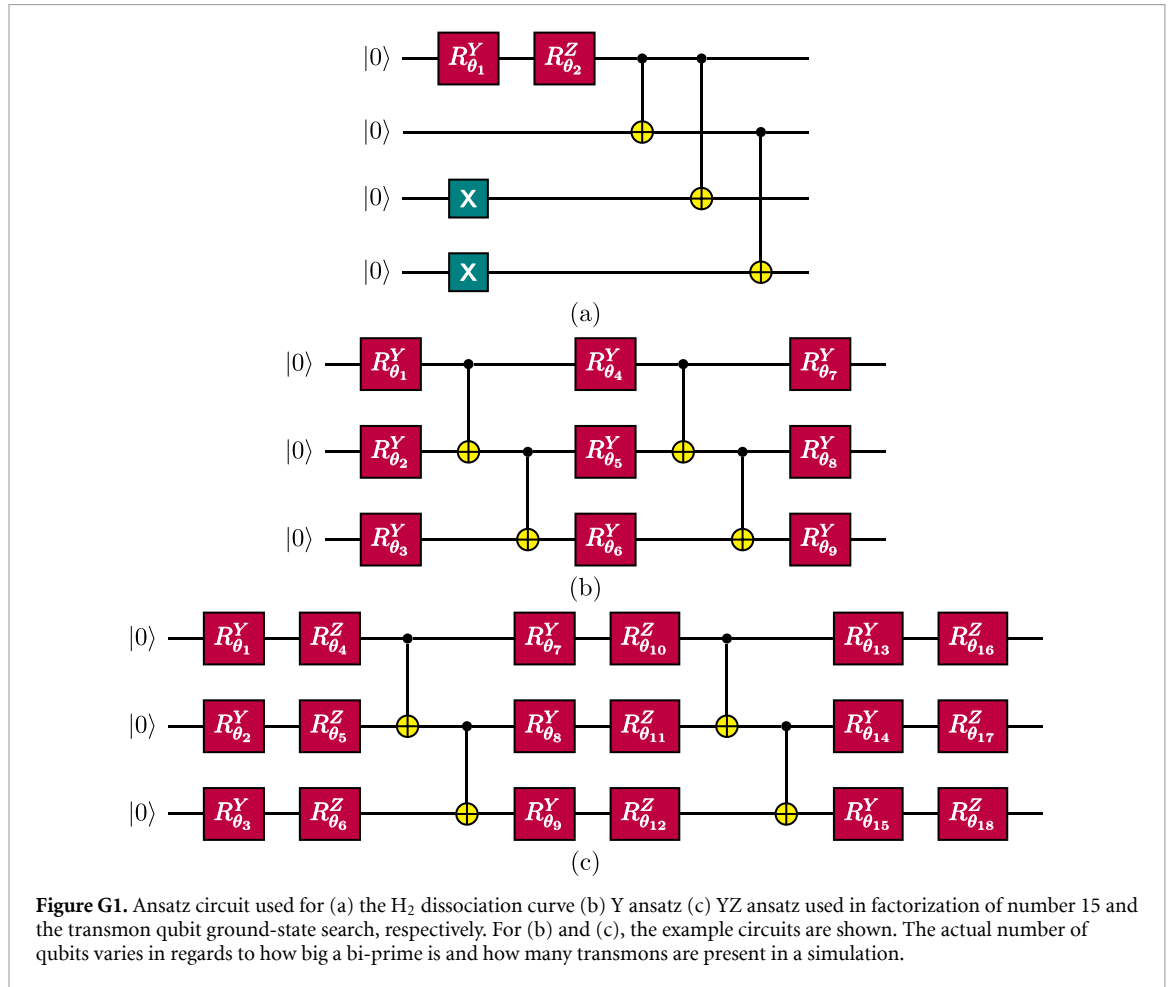
Appendix H. Molecular ground-state search

The Hamiltonian for a chemical system in the second quantization picture has the following general form

$$\hat{H} = \sum_{ij} h_{ij}\hat{a}_i^\dagger \hat{a}_j + \sum_{ijkl} V_{ijkl}\hat{a}_i^\dagger \hat{a}_k^\dagger \hat{a}_l \hat{a}_j, \tag{H.1}$$

where $\hat{a}_i^\dagger$ is a creation operator that creates an electron on the ith orbital, $\hat{a}_i$ is an annihilation operator which removes an electron from the ith orbital, and h_{ij} and V_{ijkl} are the one-electron and two-electron interaction coefficients, respectively, which are determined for specific systems. The antisymmetric property of electrons is fulfilled by the anti-commutation relation of the creation and annihilation operators $\{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij}, \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0$. The above anti-commutation relation precludes direct encoding of a chemical Hamiltonian on a quantum computer, since the operating units of a quantum computer (i.e. qubits) obey the commutation relation of spins. The remedy to this discrepancy is to perform a fermion-spin mapping, such as the Jordan–Wigner transformation [54] presented here as an example. The transformation maps the fermionic creation and annihilation operators to qubit raising and lowering operators $\hat{\sigma}^\pm = \hat{X} \pm i\hat{Y}$ with a string of $\hat{Z}$ operators to enforce the fermionic anti-commutation properties: $\hat{a}_j \rightarrow (\prod_{l=1}^{j-1} -\hat{Z}_l)\hat{\sigma}_j^-$, $\hat{a}_j^\dagger \rightarrow (\prod_{l=1}^{j-1} -\hat{Z}_l)\hat{\sigma}_j^+$ With the above transformation, the fermionic anti-commutation relation is preserved. For other fermion-spin mapping approaches, the reader is referred to the literature [4]. An alternative method to decompose gates for molecular systems is provided in [55]. Here the atomic orbital basis (sto3g basis set) is used to ground-state energy is determined at the Hartree–Fock level of theory via PySCF [56].

Appendix I. Quantum computer-aided designs

As the number of high-quality qubits inside a quantum processing unit (QPU) grows over time, it is expected that eventually no classical supercomputer will be able to simulate, verify, and cross-check the inner working mechanism and data obtained from the QPU. This is commonly taken to be as crossing over the 'quantum advantage' threshold. Once such an event occurs, from a practical point-of-view, it is beneficial to make use of existing quantum hardware that is already well-calibrated to simulate subsets of new QPU designs. QCADs of superconducting qubits [33] and photonic chips [34] have recently been proposed and experimentally realized in a superconducting qubit architecture [35], but not yet with imaginary-time-like quantum simulation. Here, we show that, with the proposed QIPA, we are able to optimize for the ground-state-energy search of a flux-tunable superconducting transmon system.

Given an arbitrary classical electrical circuit diagram composed of inductors, capacitors, and Josephson junctions, one can quantize such circuit into a quantum Hamiltonian [57] via the Legendre transformation. Once we obtain the quantum Hamiltonian, the task is to translate it into a language that a quantum computer can understand, such as Pauli words or strings. Let us consider the case of a flux-tunable transmon system shown in the main text figure 1 prior to conversion as an example. The system has the following Hamiltonian:

$$\hat{H}_{\text{transmon}} = 2e^2 \hat{N}^2 / \mathbf{C}_{\text{cap}} - 2E_J |\cos(2\pi f)| \cos\hat{\varphi}. \tag{I.1}$$

Here, e is the electron charge. The normalized external flux $f = \Phi_{\text{ext}}/\Phi_0$ is derived from the external magnetic flux Φ_{ext} that penetrates the loop formed by the two Josephson junctions of the transmon. The Josephson energy of the two junctions is equal and given by E_J while $\mathbf{C}_{\text{cap}}$ is the capacitance. The magnetic flux quantum Φ_0 is a fundamental constant that describes the smallest amount of flux that a superconducting loop can sustain. Here, $\hat{\varphi}$ and $\hat{N}$ are the phase and number operators, respectively and fulfill the commutation relation $[\hat{\varphi}, \hat{N}] = i$. And, the following relations follow:

$$[e^{i\hat{\varphi}}, \hat{N}] = -e^{i\hat{\varphi}}, \quad e^{\pm i\hat{\varphi}}|n\rangle = |n \pm 1\rangle, \tag{I.2}$$

where $|n_j\rangle$ are the eigenstates of $\hat{N}$. We notice that the operators $\widehat{e^{\pm i\varphi_j}}$ are similar to the usual bosonic creation and annihilation operators, without the square root prefactor. They are, in fact, the Susskind–Glogower phase operators [58]. When we write down the transmon Hamiltonian in the charge number basis, we use equation (I.2) and assign the operators as:

$$\hat{N} = \sum_{n=0}^{d-1} \left(n - \frac{d}{2}\right)|n\rangle\langle n|, \tag{I.3}$$

$$\cos\hat{\varphi} = \frac{1}{2}\sum_{n=0}^{d-2}(|n\rangle\langle n+1| + |n+1\rangle\langle n|), \tag{I.4}$$

$$\sin\hat{\varphi} = \frac{i}{2}\sum_{n=0}^{d-2}(|n\rangle\langle n+1| - |n+1\rangle\langle n|). \tag{I.5}$$

In general, the number of Cooper pairs can take on infinitely many integer values. However, for practical purposes, we are only interested in low-lying energy states. In that case, we can truncate the Hilbert space as described above by introducing a finite maximum number of excitations $d = 2^k$. The number of data qubits used in the quantum simulation of the transmon qubit is $k \in \mathbb{N}$.

Next, we convert the eigenstates of the number operator $\hat{N}$ into the computational basis states of k data qubits by representing the integer charge number in a preferred encoding [59–61]. This implies a truncation of the physical space to the subspace spanned by 2^k Cooper pair numbers. There are combinatorially many ways to map such a state space to a set of qubits. For all the numerical quantum simulation experiments presented in this work, we have employed the Gray code [33, 34, 61] due to its resource-efficient representation of tridiagonal quantum matrix operators (table I1).

Appendix J. Biprime factorization

Prime factorization of biprimes is essential to modern Rivest–Shamir–Adleman (RSA) encryption algorithms [62] and is seen as a classic test of the power of quantum computing to address problems that are computationally intensive on classical computers [63–66]. Prime factorization algorithms for NISQ quantum computers are essential to demonstrate the promise of quantum computers for this task and

Table I1. Qubit encodings (standard binary and Gray code) of elementary operators used in this study, with a truncation of $d = 16$. In our numerical experiments, we utilize the Gray code [61]. Here, X, Y, Z correspond to standard Pauli matrices and the subscript represents the qubit index number.

$d = 16$	Std. Binary	Gray
N	$-0.5\,I$	$-0.5\,I$
	$-4.0\,Z_3$	$-4.0\,Z_3$
	$-2.0\,Z_2$	$-2.0\,Z_2 Z_3$
	$-1.0\,Z_1$	$-1.0\,Z_1 Z_2 Z_3$
	$-0.5\,Z_0$	$-0.5\,Z_0 Z_1 Z_2 Z_3$
$\cos\varphi$	$+0.5\,X_0$	$+0.5\,X_0$
	$+0.25\,X_0 X_1$	$+0.25\,X_1$
	$+0.25\,Y_0 Y_1$	$-0.25\,Z_0 X_1$
	$+0.125\,X_0 X_1 X_2$	$+0.125\,X_2$
	$+0.125\,X_0 Y_1 Y_2$	$-0.125\,Z_1 X_2$
	$+0.125\,Y_0 X_1 Y_2$	$+0.125\,Z_0 X_2$
	$-0.125\,Y_0 Y_1 X_2$	$-0.125\,Z_0 Z_1 X_2$
	$+0.0625\,X_0 X_1 X_2 X_3$	$+0.0625\,X_3$
	$+0.0625\,X_0 X_1 Y_2 Y_3$	$-0.0625\,Z_2 X_3$
	$+0.0625\,X_0 Y_1 X_2 Y_3$	$+0.0625\,Z_1 X_3$
	$-0.0625\,X_0 Y_1 Y_2 X_3$	$-0.0625\,Z_1 Z_2 X_3$
	$+0.0625\,Y_0 X_1 X_2 Y_3$	$+0.0625\,Z_0 X_3$
	$-0.0625\,Y_0 X_1 Y_2 X_3$	$-0.0625\,Z_0 Z_2 X_3$
	$-0.0625\,Y_0 Y_1 X_2 X_3$	$+0.0625\,Z_0 Z_1 X_3$
	$-0.0625\,Y_0 Y_1 Y_2 Y_3$	$-0.0625\,Z_0 Z_1 Z_2 X_3$
$\sin\varphi$	$-0.5\,Y_0$	$-0.5\,Y_0 Z_1 Z_2 Z_3$
	$-0.25\,X_0 Y_1$	$-0.25\,Y_1 Z_2 Z_3$
	$+0.25\,Y_0 X_1$	$+0.25\,Z_0 Y_1 Z_2 Z_3$
	$-0.125\,X_0 X_1 Y_2$	$-0.125\,Y_2 Z_3$
	$+0.125\,X_0 Y_1 X_2$	$+0.125\,Z_1 Y_2 Z_3$
	$+0.125\,Y_0 X_1 X_2$	$-0.125\,Z_0 Y_2 Z_3$
	$+0.125\,Y_0 Y_1 Y_2$	$+0.125\,Z_0 Z_1 Y_2 Z_3$
	$-0.0625\,X_0 X_1 X_2 Y_3$	$-0.0625\,Y_3$
	$+0.0625\,X_0 X_1 Y_2 X_3$	$+0.0625\,Z_2 Y_3$
	$+0.0625\,X_0 Y_1 X_2 X_3$	$-0.0625\,Z_1 Y_3$
	$+0.0625\,X_0 Y_1 Y_2 Y_3$	$+0.0625\,Z_1 Z_2 Y_3$
	$+0.0625\,Y_0 X_1 X_2 X_3$	$-0.0625\,Z_0 Y_3$
	$+0.0625\,Y_0 X_1 Y_2 Y_3$	$+0.0625\,Z_0 Z_2 Y_3$
	$+0.0625\,Y_0 Y_1 X_2 Y_3$	$-0.0625\,Z_0 Z_1 Y_3$
	$-0.0625\,Y_0 Y_1 Y_2 X_3$	$+0.0625\,Z_0 Z_1 Z_2 Y_3$

complement current approaches such as Shor's algorithm [1, 63], variational quantum factoring (VQF) [67, 68], exact search [27, 65, 69], QITE [18–20, 66, 70–74], and quantum annealing [64, 75–80]. Recently QITE has been used to identify prime factors via global optimization [66], and here we show such an imaginary-time-like approach can be further accelerated by using QIPA.

To solve the prime factorization problem for a given biprime (product of two prime numbers) $N = q^* \times p^*$, we consider the Hamiltonian

$$\hat{H}_N(q,p) = d(N;q,p)^2, \quad d(N;q,p) = N - q \times p, \tag{J.1}$$

defined on the space of prime numbers $q, p \leqslant \sqrt{N}$. This Hamiltonian is a non-negative function that attains its global minimum $\hat{H}_N(q,p) = 0$ for the unique pair of solutions $q = q^*$ and $p = p^*$. Binary representation [33, 34, 61, 64, 66, 67] of the prime factors yields

$$d(N;q,p) = N - \sum_{j,k=0}^{L} q_j p_k 2^{j+k}, \tag{J.2}$$

where $(q_L, \ldots, q_0), (p_L, \ldots, p_0) \in \{0,1\}^L$ are the binary representations of q and p, respectively, and $L = \lfloor \log_2(N/2) \rfloor + 1$. Assuming that $p^*, q^* > 2$ (otherwise N is an even number), we further restrict the search space by setting $q_0 = p_0 = 1$ such that (J.2) can also be written in terms of the combined parameter
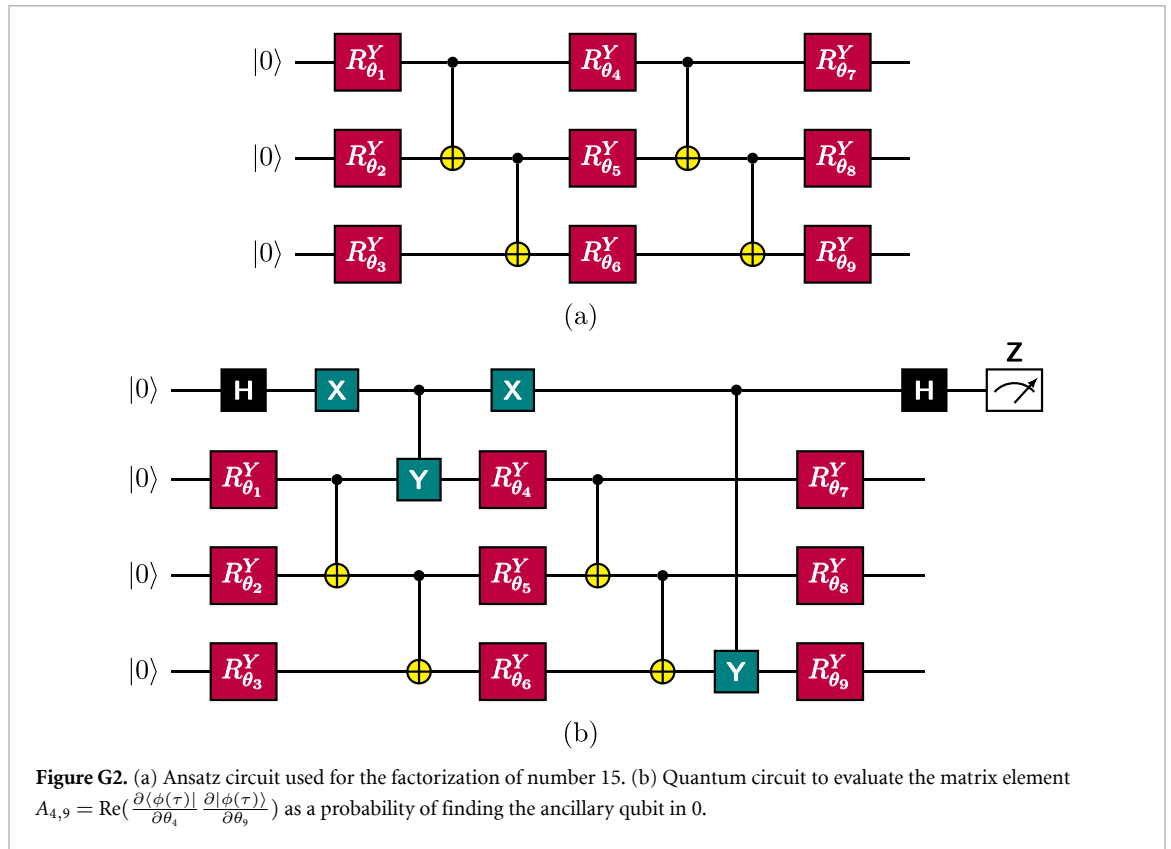
Figure G2. (a) Ansatz circuit used for the factorization of number 15. (b) Quantum circuit to evaluate the matrix element $A_{4,9} = \mathrm{Re}\left(\frac{\partial\langle\phi(\tau)|}{\partial\theta_4}\frac{\partial|\phi(\tau)\rangle}{\partial\theta_9}\right)$ as a probability of finding the ancillary qubit in 0.

$\vec{x} = (q_1,\ldots,q_L,p_1,\ldots,p_L)$, as follows:

$$d(N;\vec{x}) = N - \left(1 + \sum_{j=1}^{L} x_j 2^j\right) \times \left(1 + \sum_{k=1}^{L} x_{L+k} 2^k\right), \tag{J.3}$$

where $x_l^2 = x_l$ holds for all $1 \leqslant l \leqslant 2L$ since $x_l \in \{0,1\}$. Equivalently, in terms of the scaled spin parameters $s_l = 2x_l - 1$, we write

$$d(N;\vec{s}) = N - \left(2^L + \sum_{j=1}^{L} s_j 2^{j-1}\right) \times \ldots \left(2^L + \sum_{k=1}^{L} s_{L+k} 2^{k-1}\right), \tag{J.4}$$

where $s_l^2 = 1$ holds since $s_l \in \{-1,1\}$. For example, for $N = 15$ we obtain $L = 2$ and $d(15;\vec{s}) = 15 - (4 + s_1 + 2s_2) \times (4 + s_3 + 2s_4) = -1 - 4s_1 - 8s_2 - 4s_3 - s_1 s_3 - 2s_2 s_3 - 8s_4 - 2s_1 s_4 - 4s_2 s_4$, which gives the Hamiltonian $H_{15}(\vec{s}) = d(15;\vec{s})^2 = 186 + 48s_1 + 96s_2 + 84s_1 s_2 + 48s_3 + 34s_1 s_3 + 68s_2 s_3 + 32s_1 s_2 s_3 + 96s_4 + 68s_1 s_4 + 136s_2 s_4 + 64s_1 s_2 s_4 + 84s_3 s_4 + 32s_1 s_3 s_4 + 64s_2 s_3 s_4 + 16s_1 s_2 s_3 s_4$, with twin global minima. Alternatively, with previous information about the prime factors of 15, the integer can be factorized with fewer qubits via the test Hamiltonian [66] $\hat{H} = 196 - 52\hat{Z}_2 - 52\hat{Z}_0 - 56\hat{Z}_2\hat{Z}_0 - 96\hat{Z}_1 - 48\hat{Z}_2\hat{Z}_1 + 16\hat{Z}_0\hat{Z}_1 + 128\hat{Z}_0\hat{Z}_1\hat{Z}_2$, which features a unique ground state $|011\rangle$ that corresponds to the correct factorization of the number 15 to 3 and 5. The resulting circuit implementation is given in figure G2.

ORCID iDs

Thi Ha Kyaw ⦿ https://orcid.org/0000-0002-3557-2709
Micheline B Soley ⦿ https://orcid.org/0000-0001-7973-2842
Brandon Allen ⦿ https://orcid.org/0000-0002-5512-1892
Paul Bergold ⦿ https://orcid.org/0000-0003-3033-0732
Chong Sun ⦿ https://orcid.org/0000-0002-8299-9094
Victor S Batista ⦿ https://orcid.org/0000-0002-3262-1237
Alán Aspuru-Guzik ⦿ https://orcid.org/0000-0002-8277-4434

References

[1] Shor P W 1994 Algorithms for quantum computation: discrete logarithms and factoring *Proc. 35th Annual Symp. on Foundations of Computer Science* (IEEE) pp 124–34

[2] Aspuru-Guzik A, Dutoi A D, Love P J and Head-Gordon M 2005 *Science* **309** 1704–7

[3] Georgescu I M, Ashhab S and Nori F 2014 *Rev. Mod. Phys.* **86** 153–85

[4] Cao Y *et al* 2019 *Chem. Rev.* **119** 10856–915

[5] McArdle S, Endo S, Aspuru-Guzik A, Benjamin S C and Yuan X 2020 *Rev. Mod. Phys.* **92** 015003

[6] Preskill J 2018 *Quantum* **2** 79

[7] Arute F *et al* 2019 *Nature* **574** 505–10

[8] Zhong H S *et al* 2020 *Science* **370** 1460–3

[9] Wu Y *et al* 2021 *Phys. Rev. Lett.* **127** 180501

[10] Madsen L S *et al* 2022 *Nature* **606** 75–81

[11] Bharti K *et al* 2022 *Rev. Mod. Phys.* **94** 015004

[12] Cerezo M *et al* 2021 *Nat. Rev. Phys.* **3** 625–44

[13] Tilly J *et al* 2021 arXiv:2111.05176

[14] Fedorov D A, Peng B, Govind N and Alexeev Y 2022 *Mater. Theory* **6** 1–21

[15] Peruzzo A, McClean J, Shadbolt P, Yung M-H, Zhou X-Q, Love P J, Aspuru-Guzik A and O'Brien J L 2014 *Nat. Commun.* **5** 1–7

[16] McClean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 *New J. Phys.* **18** 023023

[17] Li Y and Benjamin S C 2017 *Phys. Rev.* X **7** 021050

[18] McArdle S, Jones T, Endo S, Li Y, Benjamin S C and Yuan X 2019 *npj Quantum Inf.* **5** 1–6

[19] Yuan X, Endo S, Zhao Q, Li Y and Benjamin S C 2019 *Quantum* **3** 191

[20] Motta M, Sun C, Tan A T K, O'Rourke M J, Ye E, Minnich A J, Brandão F G S L and Chan G K L 2020 *Nat. Phys.* **16** 205–10

[21] Bottou L, Curtis F E and Nocedal J 2018 *SIAM Rev.* **60** 223–311

[22] Wales D J and Scheraga H A 1999 *Science* **285** 1368–72

[23] Dill K A, Ozkan S B, Shell M S and Weikl T R 2008 *Annu. Rev. Biophys.* **37** 289

[24] Coutinho W P, Battarra M and Fliege J 2018 *Comput. Ind. Eng.* **120** 116–28

[25] Levinthal C 1969 How to fold graciously *Mossbauer Spectroscopy in Biological Systems Proceedings* vol 67 pp 22–24

[26] Venter G 2010 *Review of Optimization Techniques* (Wiley)

[27] Grover L K 1996 A fast quantum mechanical algorithm for database search *STOC '96: Proc. 28th Annual ACM Symp. on Theory of Computing* (Association for Computing Machinery) pp 212–9

[28] Oseledets I and Tyrtyshnikov E 2010 *Linear Algebra Appl.* **432** 70–88

[29] Oseledets I V 2011 *SIAM J. Sci. Comput.* **33** 2295–317

[30] Soley M B, Bergold P and Batista V S 2021 *J. Chem. Theory Comput.* **17** 3280–91

[31] Sim S, Johnson P D and Aspuru-Guzik A 2019 *Adv. Quantum Technol.* **2** 1900070

[32] Lee S *et al* 2022 arXiv:2208.02199

[33] Kyaw T H, Menke T, Sim S, Anand A, Sawaya N P D, Oliver W D, Guerreschi G G and Aspuru-Guzik A 2021 *Phys. Rev. Appl.* **16** 044042

[34] Kottmann J, Krenn M, Kyaw T H, Alperin-Lea S and Aspuru-Guzik A 2021 *Quantum Sci. Technol.* **6** 035010

[35] Liu F M *et al* 2021 arXiv:2109.00994

[36] Endo S, Cai Z, Benjamin S C and Yuan X 2021 *J. Phys. Soc. Japan* **90** 032001

[37] Endo S, Benjamin S C and Li Y 2018 *Phys. Rev.* X **8** 031027

[38] Temme K, Bravyi S and Gambetta J M 2017 *Phys. Rev. Lett.* **119** 180509

[39] Otten M and Gray S K 2019 *npj Quantum Inf.* **5** 1–6

[40] Strikis A, Qin D, Chen Y, Benjamin S C and Li Y 2020 arXiv:2005.07601

[41] Bravyi S, Sheldon S, Kandala A, Mckay D C and Gambetta J M 2021 *Phys. Rev.* A **103** 042605

[42] Gomes N, Mukherjee A, Zhang F, Iadecola T, Wang C-Z, Ho K-M, Orth P P and Yao Y-X 2021 *Adv. Quantum Technol.* **4** 2100114

[43] Zhang Z J, Sun J, Yuan X and Yung M H 2020 arXiv:2011.05283

[44] Loken C *et al* 2010 *J. Phys.: Conf. Ser.* **256** 012026

[45] Ponce M *et al* 2019 *Proc. Practice and Experience in Advanced Research Computing on Rise of the Machines (Learning)* vol 34 p 1

[46] Kottmann J S *et al* 2021 *Quantum Sci. Technol.* **6** 024009

[47] Suzuki Y *et al* 2021 *Quantum* **5** 559

[48] McLachlan A 1964 *Mol. Phys.* **8** 39–44

[49] Broeckhove J, Lathouwers L, Kesteloot E and Van Leuven P 1988 *Chem. Phys. Lett.* **149** 547–50

[50] Poulin D, Qarry A, Somma R and Verstraete F 2011 *Phys. Rev. Lett.* **106** 170501

[51] Hestenes M R and Stiefel E 1952 *J. Res. Natl Bur. Stand.* **49** 409

[52] Romero J, Babbush R, McClean J R, Hempel C, Love P J and Aspuru-Guzik A 2018 *Quantum Sci. Technol.* **4** 014008

[53] Ekert A K, Alves C M, Oi D K, Horodecki M, Horodecki P and Kwek L C 2002 *Phys. Rev. Lett.* **88** 217901

[54] Jordan P and Wigner E 1928 *Z. Phys.* **47** 631–51

[55] Daskin A and Kais S 2011 *J. Chem. Phys.* **134** 144112

[56] Sun Q *et al* 2018 *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **8** e1340

[57] Kyaw T H 2019 *Towards a Scalable Quantum Computing Platform in the Ultrastrong Coupling Regime* (Springer)

[58] Susskind L and Glogower J 1964 *Phys. Phys. Fiz.* **1** 49

[59] Veis L, Višňák J, Nishizawa H, Nakai H and Pittner J 2016 *Int. J. Quantum Chem.* **116** 1328

[60] McArdle S, Mayorov A, Shan X, Benjamin S and Yuan X 2019 *Chem. Sci.* **10** 5725

[61] Sawaya N P D, Menke T, Kyaw T H, Johri S, Aspuru-Guzik A and Guerreschi G G 2020 *npj Quantum Inf.* **6** 49

[62] Rivest R L, Shamir A and Adleman L 1978 *Commun. ACM* **21** 120–6

[63] Martín-López E, Laing A, Lawson T, Alvarez R, Zhou X-Q and O'Brien J L 2012 *Nat. Photon.* **6** 773–6

[64] Jiang S, Britt K A, McCaskey A J, Humble T S and Kais S 2018 *Sci. Rep.* **8** 17667

[65] Dash A, Sarmah D, Behera B K and Panigrahi P K 2018 arXiv:1805.10478v2

[66] Selvarajan R, Dixit V, Cui X, Humble T S and Kais S 2021 *Sci. Rep.* **11** 1–8

[67] Anschuetz E R, Olson J P, Aspuru-Guzik A and Cao Y 2018 arXiv:1808.08927v1

[68] Karamlou A H, Simon W A, Katabarwa A, Scholten T L, Peropadre B and Cao Y 2021 *npj Quantum Inf.* **7** 1–6

[69] Liu Y 2014 *Int. J. Theor. Phys.* **53** 2571–8
[70] Moll N *et al* 2018 *Quantum Sci. Technol.* **3** 030503
[71] Gomes N, Zhang F, Berthusen N F, Wang C-Z, Ho K-M, Orth P P and Yao Y 2020 *J. Chem. Theory Comput.* **16** 6256–66
[72] Nishi H, Kosugi T and Matsushita Y I 2021 *npj Quantum Inf.* **7** 1–7
[73] Huang Y, Shao Y, Ren W, Sun J and Lv D 2022 arXiv:2203.11112
[74] Yeter-Aydeniz K, Moschandreou E and Siopsis G 2022 *Phys. Rev.* A **105** 012412
[75] Kadowaki T and Nishimori H 1998 *Phys. Rev.* E **58** 5355–63
[76] Farhi E, Goldstone J, Gutmann S and Sipser M 2000 arXiv:quant-ph/0001106
[77] Farhi E, Goldstone J, Gutmann S, Lapan J, Lundgren A and Preda D 2001 *Science* **292** 472–5
[78] Perdomo-Ortiz A, Dickson N, Drew-Brook M, Rose G and Aspuru-Guzik A 2012 *Sci. Rep.* **2** 1–7
[79] Babbush R, Love P J and Aspuru-Guzik A 2014 *Sci. Rep.* **4** 1–11
[80] Albash T and Lidar D A 2018 *Phys. Rev.* X **8** 031016