

# Rapid initial state preparation for the quantum simulation of strongly correlated molecules

Dominic W. Berry,<sup>1,\*</sup> Yu Tong,<sup>2,3,4</sup> Tanuj Khattar,<sup>5</sup> Alec White,<sup>5</sup> Tae In Kim,<sup>6</sup> Sergio Boixo,<sup>5</sup> Lin Lin,<sup>7</sup> Seunghoon Lee,<sup>6,8</sup> Garnet Kin-Lic Chan,<sup>8</sup> Ryan Babbush,<sup>5</sup> and Nicholas C. Rubin<sup>5,†</sup>

<sup>1</sup>*School of Mathematical and Physical Sciences, Macquarie University, Sydney, NSW 2109, Australia*

<sup>2</sup>*Institute of Quantum Information and Matter, California Institute of Technology, Pasadena, CA 91125, United States*

<sup>3</sup>*Department of Mathematics, Duke University, Durham, NC 27708, United States*

<sup>4</sup>*Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708, United States*

<sup>5</sup>*Google Quantum AI, Venice, CA 90291, United States*

<sup>6</sup>*Department of Chemistry, Seoul National University, Seoul 08826, Republic of Korea*

<sup>7</sup>*Department of Mathematics, University of California, Berkeley, CA 94720, United States*

<sup>8</sup>*Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, United States*

(Dated: September 19, 2024)

Studies on quantum algorithms for ground state energy estimation often assume perfect ground state preparation; however, in reality the initial state will have imperfect overlap with the true ground state. Here we address that problem in two ways: by faster preparation of matrix product state (MPS) approximations, and more efficient filtering of the prepared state to find the ground state energy. We show how to achieve unitary synthesis with a Toffoli complexity about  $7\times$  lower than that in prior work, and use that to derive a more efficient MPS preparation method. For filtering we present two different approaches: sampling and binary search. For both we use the theory of window functions to avoid large phase errors and minimise the complexity. We find that the binary search approach provides better scaling with the overlap at the cost of a larger constant factor, such that it will be preferred for overlaps less than about 0.003. Finally, we estimate the total resources to perform ground state energy estimation of Fe-S cluster systems, including the FeMo cofactor by estimating the overlap of different MPS initial states with potential ground-states of the FeMo cofactor using an extrapolation procedure. With a modest MPS bond dimension of 4000, our procedure produces an estimate of  $\sim 0.9$  overlap squared with a candidate ground-state of the FeMo cofactor, producing a total resource estimate of  $7.3 \times 10^{10}$  Toffoli gates; neglecting the search over candidates and assuming the accuracy of the extrapolation, this validates prior estimates that used perfect ground state overlap. This presents an example of a practical path to prepare states of high overlap in a challenging-to-compute chemical system.

## CONTENTS

I. Introduction	2
A. Results overview	4
II. Background	5
A. Matrix product state preparation	5
B. Phase estimation	6
1. The quantum phase estimation circuit	7
2. The phase error	7
3. Window functions	8
III. MPS preparation	9
A. New unitary synthesis method	9
B. Synthesis of columns of unitary	12
IV. Phase estimation with window functions	14
A. The Kaiser window	14
B. Higher-order terms	15
C. The prolate spheroidal window	17

---

\* dominic.berry@mq.edu.au

† nickrubin@google.com

V. The sampling method	18
A. Asymptotic approximations	18
B. Real cost for window functions	21
C. Contribution to cost from excited states	21
VI. The binary search approach	23
A. From the fuzzy bisection problem to amplitude estimation	23
B. Amplitude estimation with the Kaiser window	25
C. Query complexity of all search steps	26
D. Optimizing the shrinking factor	26
VII. Total resources for ground state energy estimation in chemical systems	28
A. Model generation	29
B. Overlap extrapolation protocol	30
C. MPS preparation costs versus phase estimation costs	32
VIII. Conclusions	33
Acknowledgements	34
References	34
A. Further details of higher-order approximations	35
B. Slepian expansion	37
C. Methods of calculating prolate spheroidal functions	39
D. Combinations of excited states	40
E. Costing for LKS scheme	41
F. Threshold analysis for resource estimates	42
1. $\text{Fe}_2\text{S}_2$	42
2. $\text{Fe}_4\text{S}_4$	43
3. $\text{FeMoCo}$	46
G. Validation of overlap extrapolation protocol	46

## I. INTRODUCTION

The complexity of estimating ground state energies of chemical and material systems using quantum phase estimation (QPE) is frequently analysed in the ideal case where the ground state has already been prepared accurately. In this restrictive setting, the main error in QPE originates from the eigenphase differing from the approximate output from QPE—a phenomenon known as “spectral leakage” [1] or “bit discretization error” [2]. In the case where the ground state (or more generally an eigenstate) is not prepared exactly, the QPE protocol outputs an estimate of the ground state energy with a probability proportional to the square of the overlap of the ground state and the input initial state. In order to estimate the total amount of quantum resources (logical qubits and gates) for the most important simulation problems and determine total runtimes for high confidence eigenenergies, both sources of error in QPE must be quantified and accounted for, along with the cost of performing phase estimation.

The problem of preparing an initial state with high overlap with the ground state has been of recent interest. This is in part because a generic state in the Hilbert space has exponentially small overlap with the ground state, and a non-trivial state preparation might need to be performed for a reasonable success probability in QPE, which could add substantially to the total QPE cost. This has spurred a substantial body of work examining the cost of preparing various approximate wavefunctions and the associated success probability of QPE. For example, analysis of product state wavefunctions in the fermionic setting [3] along with informed orbital optimized improvements [4], state preparations analyzed in the context of embedding theories [5], truncated configuration interaction [6], and the use of matrix-product states (MPS) as input QPE states [7]. In terms of circuit compilations, direct synthesis costs for wavefunctions are known [8] along with a variety of MPS state preparation techniques, including layers of two-qubit

operations [9], low-depth circuits [10], and sequences of operations with an ancilla register [11]. Reference [11] in particular has been analysed in detail for minimising the Toffoli count [6]. A related consideration in the context of evaluating quantum advantage is that classical heuristic algorithms can also be viewed as having a state preparation step, from which a classical estimate of the ground-state energy is then estimated and, in some cases, efficiently refined [12, 13]. Hence it is relevant to ask for specific systems, the quantitative cost of preparing a good initial state for QPE.

Given that an appropriate initial state has been prepared, there remains the problem of how to measure the ground state energy in a way that properly distinguishes it from excited states. If the overlap is not too small, then it may be expected that a simple sampling approach will suffice, and it will be necessary to sample enough times that there will be a high probability of sampling the ground state energy at least once. The sample that corresponds to the ground state energy can be identified by taking the minimum among all samples. There are two difficulties with this approach. First, the large number of samples may mean that there is exceptionally large underestimation error in at least one of the samples, resulting in an erroneously small estimate of the ground state energy. Second, the number of samples will scale with the inverse of the squared overlap.

Amplitude amplification would suggest that the complexity should scale with the inverse overlap instead of the inverse overlap squared. Naively, in order to obtain the quadratic improvement, the range of energies for the ground state must be known. Binary search can be employed to go beyond this requirement and search for an unknown ground state. The binary search approach was used previously in [14–18], but these works, together with [19–21], typically aim to optimize for the circuit depth, while this work focuses on the number of non-Clifford gates needed for the whole algorithm, a more relevant metric for fault-tolerant quantum computers.

In this work we provide improved results for both MPS preparation and filtering to determine the ground state energy. For MPS preparation we start by developing a method for synthesising unitaries with low Toffoli count by decomposing the unitary into a sequence of diagonal phasing operations together with low-cost operations. We then use that to construct a method for synthesising only a fraction of the columns of the unitary, which we then apply to the method of Ref. [11] to provide a factor of 7 improvement in Toffoli gates over prior work. For filtering the resulting state, we apply the theory of window functions in order to minimise the probability of estimates with large error. We describe the optimal performance provided by the Slepian prolate spheroidal window [22–24] and compare the costs to the Kaiser window previously reported in Ref. [25]. In the case of sampling, this suppresses the probability of exceptionally low energy estimates that would make the overall estimate low, overcoming the first problem. Moreover, we provide a tighter bound on the contribution to the error from excited states, by analysing the interplay between the contribution to errors where the estimated energy is too low versus too high. We find that Kaiser windows can provide substantially improved performance.

We also provide improved scaling with the overlap by using a binary search together with amplitude estimation, similar to Ref. [26]. We successively reduce the possible range for the ground state by performing amplitude estimation at each step in order to eliminate a fraction of the range. In this way we are able to achieve the speedup promised by amplitude amplification without any initial estimate of the ground state energy. On the other hand, the overhead induced by this procedure means that it is preferable for small overlaps  $p \lesssim 0.003$ , and for large overlaps the sampling method is preferable. We apply the optimal windows for phase estimation to the amplitude estimation procedure, and thereby significantly reduce the resources compared to Ref. [26].

Armed with a detailed costing of the number of times QPE must be repeated, we estimate the full quantum resources necessary to refine the ground state energy of several Fe-S clusters: [2Fe-2S], [4Fe-4S], and FeMo-cofactor (FeMoco). In order to obtain overlaps of low bond dimension MPS states with the true ground state we introduce an extrapolation protocol that uses two MPS wavefunctions to derive an empirical estimate of the overlap of a fixed bond dimension wavefunction and the infinite (exact) bond dimension MPS. In FeMoco, at a finite bond dimension we can obtain low-bond dimension MPS that are candidates for different low-energy states of the cluster, although these initial MPS do not give a reliable energy ordering. For each candidate MPS, we can estimate the overlap with the eventual eigenstate, which allows us to cost out energy estimation for FeMoco in the high-confidence regime (95% and 99% confidence level) to chemical accuracy. We also re-analyze the block encoding costs for FeMoco and other Fe-S clusters using symmetry shifting, resulting in a reduction of the LCU 1-norm by a factor of up to 2. We find that very few iterations of QPE are required (2 or 3 iterations) due to the high overlap of low-bond dimension MPS. These additional iterations over single-shot QPE resource estimates provided for FeMoco in Ref. [27] along with symmetry shifting reductions in the LCU 1-norm result in  $7.3 \times 10^{10}$  Toffoli gates required for a full resource cost to refine an energy estimate for FeMoco. For a single candidate ground state, this amount of resources is only 2.3 times that in Ref. [27] (which uses the Hamiltonian defined in Ref. [28]) and can likely be reduced further through improved symmetry shifting. Once the accurate energies of different candidates are obtained, they can be ordered to determine the ground-state energy. Ultimately, the high extrapolated overlap achieved in this problem suggests that the combination of generating one (or more) candidate MPS initial states followed by QPE is a practical approach to refine the ground-state energy in a realistic challenging chemical system.

In the following we begin in Section IA by summarising the results that will be presented. We then give the background for both MPS preparation and phase estimation in Section II. Section III describes our new method for unitary synthesis and preparing MPS states. Then Section IV summarises the procedure to perform phase estimation optimally for confidence intervals. This is a procedure which is used for both approaches to searching for the ground state energy. In Section V we describe the sampling approach, and in Section VI we describe the binary search approach. These results are used to estimate resource requirements for real systems in Section VII, and we conclude in Section VIII.

### A. Results overview

We provide results for three main areas: preparation of matrix product states, energy estimation with these prepared states, and resource estimates for ground state energy estimates of the FeMoco chemical system.

For MPS preparation, we first derive a new result for the synthesis of general unitary operations with reduced Toffoli count. This method reduces the synthesis to layers of phase shifts alternating with increment/decrement operations and Hadamard gates. The layers of phase shifts can be applied with reduced Toffoli count using QROM. As a result there are at most  $N_{\text{un}} + 1$  layers of QROM for dimension  $N_{\text{un}}$ , and each QROM has complexity  $\mathcal{O}(\sqrt{N_{\text{un}}})$ .

We then use this result to derive a procedure to synthesise columns of a unitary operation. That is, a unitary operation where the input state is restricted to lie within a subspace. If only half the columns of the unitary need be synthesised, we reduce the problem to synthesis of an initial unitary of dimension  $N_{\text{un}}/2$ , a controlled qubit rotation, then controlled synthesis of a unitary of dimension  $N_{\text{un}}/2$ . That controlled unitary has half the complexity of synthesising a unitary operation of dimension  $N_{\text{un}}$ , because only half the layers are required. Then in order to synthesise  $N_{\text{un}}/d$  columns of a unitary we can iterate this procedure another  $d - 2$  times.

The MPS preparation can then be achieved by a sequence of steps where  $N_{\text{un}}/d$  columns of a unitary need to be synthesised. Moreover, the initial unitary in the above procedure can be merged with other operations, further reducing the complexity. As a result, the overall Toffoli complexity is significantly reduced over that obtained for methods in prior work, now making it a small complexity as compared to the complexity of phase estimation.

For energy estimation, there are two approaches we consider: direct sampling, and a binary search with amplitude estimation. The parameters of the problem are

- the initial squared overlap of the prepared state with the ground state,  $p$ ,
- the block encoding normalization factor  $\lambda$ ,
- the allowable error in the energy estimate  $\epsilon$ , and
- the confidence level  $1 - q$ .

For the direct sampling approach, we provide expressions to determine exact costings in terms of special functions. We then derive asymptotic expressions to provide the expected scaling of the complexity in the above parameters. The total number of queries to the qubitized walk operator encoding the Hamiltonian is approximately

$$\frac{n\lambda}{2\epsilon} \ln(1/\delta) \approx \frac{\lambda \ln(2/q)}{2p\epsilon} \ln \left[ \frac{\ln(2/q)}{pq} \right], \quad (1)$$

corresponding to the leading term in Eq. (85). This expression is found by solving

$$[1 - p(1 - \delta/2)]^n + 1 - (1 - \delta/2)^n = q \quad (2)$$

for  $\delta$ , then minimising with respect to the number of samples  $n$ . The optimal choice for  $n$  is slightly larger than  $(1/p) \ln(1/q)$ .

For the binary search approach, under the same conditions, the query complexity is

$$\frac{7.77\lambda}{\sqrt{p}\epsilon} \ln \left( \frac{4}{\sqrt{p}} \right) \ln \left( \frac{\log_{\sqrt{2}}(\lambda/\epsilon)}{q} \right), \quad (3)$$

as in Eq. (127). This has improved scaling as  $1/\sqrt{p}$ , which is a square root improvement, albeit with a constant factor about 16 times larger. This constant factor suggests that we would need  $p \lesssim 0.003$  for this approach to provide an improvement.

These asymptotic results can be inaccurate for realistic values of parameters, so we provide improved approximations for phase measurements with window functions in Section IV B. We provide series expansions for the error and cost, both for the Kaiser window and for the prolate spheroidal window. In the process we correct an error in the work of Slepian from 1965 [22], which gave incorrect terms. We show that the error is asymptotically lower for the prolate spheroidal window, but the cost (the number of oracle calls to achieve a given error) is asymptotically the same for the two windows.

We show how to properly account for the excited states in sampling to obtain the ground state. These states contribute to two types of error.

Type I: The excited states contribute to the probability of a sample with large error, more than  $\epsilon$  below the ground state energy.

Type II: Samples corresponding to excited states contribute to estimates that are too high, because it is an accurate (or high) estimate of that energy, but the excited state energy is higher than the ground state energy.

When excited state energies are close to the ground state, then they increase the probability of a Type I error, but that also reduces the probability of a Type II error. We provide a careful accounting of the contribution of these two errors to provide a more accurate estimate of the cost of the sampling approach. Surprisingly, we find that the Kaiser window can provide *better* results than the prolate spheroidal window when accounting for excited states. Taking all these considerations into account yields results close to those for the approximate asymptotic expression given above.

We provide numerical results for query complexities accounting both for the exact error with window functions and the effect of the excited states. The results for queries to block encoding of the Hamiltonian are provided in Fig. 13, and for the number of calls to the initial state preparation in Fig. 14. This shows that the binary search approach is optimal for small squared overlap  $p$ , but the sampling approach is optimal for moderate to large values of  $p$  (above  $p \sim 0.003$  for the case of 95% confidence intervals).

For FeMoco resource estimates we develop an extrapolation scheme that allows us to estimate the overlap of a fixed bond dimension MPS with an infinite bond dimension MPS—e.g. the true eigenstate. The extrapolation protocol is constructed from two empirically observed linear relationships

$$\log \left( 1 - |\langle \Phi(M') | \Phi(\infty) \rangle|^2 \right) \text{ vs. } (\log(M'))^2 \quad (4)$$

$$\log \left( |\langle \Phi(M') | \Phi(M'') \rangle|^2 - |\langle \Phi(M') | \Phi(\infty) \rangle|^2 \right) \text{ vs. } (\log(M''))^2 \text{ where } M' \ll M'' \quad (5)$$

and verified on  $\text{Fe}_2\text{S}_2$  and  $\text{Fe}_4\text{S}_4$  systems where accurate estimates of the ground states can be computed, and where the above extrapolation can be verified. We analyze three different MPS wave functions for FeMoco initialized through a procedure similar to Ref. [28, 29] that is believed to generate candidates for competing low-energy eigenstates (corresponding to different spin couplings) of the  $S = 3/2$  ground state. The high overlaps estimated with some eigenstate in the low-energy manifold, as produced by our protocol, suggest that the combination of initializing different candidate MPS states, followed by QPE, is a promising computational procedure to map out the lowest eigenenergies and subsequently refine the ground-state energy in this practically challenging chemical simulation problem.

## II. BACKGROUND

### A. Matrix product state preparation

Matrix product states provide a systematically improvable approximation of entangled states thus providing a class of initial states with tunable overlap. The Toffoli complexity of MPS preparation was previously analysed in Ref. [6] using the approach from Ref. [11] together with the unitary synthesis scheme of Low, Kliuchnikov and Schaeffer (LKS) Ref. [8]. In this work we provide a significantly more efficient unitary synthesis scheme than that of Ref. [8], thereby enabling more efficient MPS preparation.

MPS states for  $\mathcal{N}$  subsystems of dimension  $d$  are of the form

$$\sum_{\{n\}} \text{Tr} \left[ A_1^{(n_1)} A_2^{(n_2)} \dots A_{\mathcal{N}}^{(n_{\mathcal{N}})} \right] |n_1, n_2, \dots, n_{\mathcal{N}}\rangle. \quad (6)$$

The matrices  $A_j^{(n_j)}$  are of dimension  $\chi$ , called the bond dimension, and the indices  $n_j$  range over  $d$  values. The principle of the approach from Ref. [11] is to use an ancilla of dimension  $\chi$ , together with a sequence of unitary operations on this ancilla together with the subsystems. Using the fact that one is free to represent the MPS of Eq. (6) in left canonical form, the matrices of the MPS can be cast as unitaries such that

$$G[j]_{\alpha_j n_j, \alpha_{j-1} 0} = (A_j^{(n_j)})_{\alpha_{j-1}, \alpha_j}. \quad (7)$$

That is, the unitaries  $G[j]$  are of dimension  $d\chi$ , but only  $\chi$  columns are specified due to the input on the physical leg being zero. The notation  $(A_j^{(n_j)})_{\alpha_{j-1}, \alpha_j}$  indicates the matrix element  $\alpha_{j-1}, \alpha_j$  of matrix  $A_j^{(n_j)}$ . There is a requirement for this technique that the specified columns are orthonormal so they may correspond to columns of a unitary.

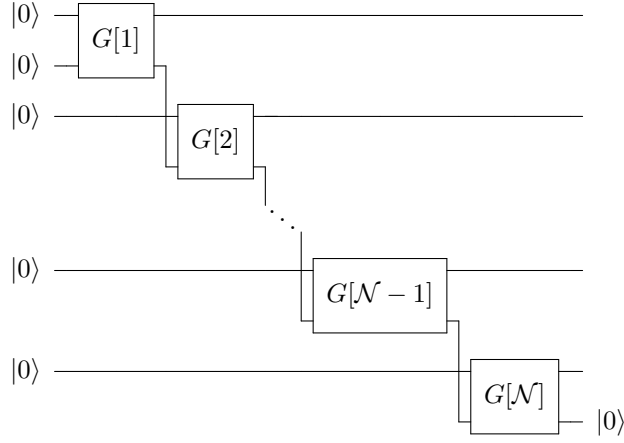


FIG. 1. The sequence of operations used to prepare a MPS. The input state  $|0\rangle$  second from the top is of dimension  $\chi$ , as is the final output state at the bottom  $|0\rangle$ .

The sequence of unitary operations used to prepare the MPS is shown in Fig. 1. There are three distinct cases where these unitary operations are performed.

1. The initial unitary  $G[1]$  is on dimension  $d\chi$ , but is guaranteed to have input states that are zeroed. Therefore it corresponds to simply preparing a state of dimension  $d\chi$ , which is simpler than synthesising a general unitary operation on this dimension.
2. There are  $\mathcal{N} - 2$  unitaries  $G[2]$  to  $G[\mathcal{N} - 1]$  on dimension  $d\chi$  where one of the input registers is initialised to  $|0\rangle$ , but the other is the ancilla which may be in a general state entangled with other registers. Therefore, only  $\chi$  columns of the unitary need be synthesised.
3. The final unitary  $G[\mathcal{N}]$  is required to reset the ancilla to  $|0\rangle$ .

For this last operation to be possible, it must be the case that the ancilla prior to the operation has support on dimension  $d$ . It must therefore be possible to transform to the required final state (with other qubits being zeroed) via a unitary without introducing an ancilla. This unitary can therefore be combined with  $G[\mathcal{N} - 1]$  to give a single unitary, and  $G[\mathcal{N}]$  does not contribute to the cost.

## B. Phase estimation

In phase estimation for ground state energy estimation, a standard approach [30, 31] is to construct a walk operator  $W$  from a block encoding of the Hamiltonian, which yields eigenvalues  $e^{\pm i\phi_j}$ , with

$$\phi_j = \arccos(\lambda_j/\lambda). \quad (8)$$

Here  $\lambda$  is the constant in the block encoding of  $H$ ; that is, the block encoding gives  $H/\lambda$ . Note that there are two conventions in the definition of  $W$  depending on what factor of  $i$  is included (corresponding to the two references [30] and [31]). The other convention yields eigenvalues  $\mp e^{\mp i \arcsin(\lambda_j/\lambda)}$ , but use of either convention yields equivalent results.

From measuring  $\pm\phi_j$  we can recover  $\lambda_j$  via  $\lambda_j = \lambda \cos(\phi_j)$ . This method has a number of useful features.

1. The cosine function is even so eliminates the  $\pm$  sign ambiguity.
2. The function  $\lambda \cos(\phi_j)$  is monotonically decreasing in  $\phi_j$ , from  $\lambda$  for  $\phi_j = 0$  to  $-\lambda$  for  $\phi_j = \pi$ . (We use the convention that  $\arccos$  gives values in the range  $[0, \pi]$ .) Therefore  $\lambda_0$ , which is the smallest among all  $\lambda_j$ , corresponds to the largest among  $\phi_j$ .
3. Because  $|\cos(x) - \cos(y)| \leq |x - y|$  for all  $x, y \in \mathbb{R}$ , if we want to estimate  $\lambda_0$  to within additive error  $\epsilon$ , it suffices to estimate  $\phi_0$  to within  $\epsilon/\lambda$ .

Each time we run the quantum phase estimation circuit, we will get a phase estimate  $\hat{\phi}$  that corresponds to  $\phi_j$  for some  $j$ , up to a phase error that we will define later. Each phase estimate then gives us an estimate for the corresponding eigenvalue  $\lambda_j$  through

$$\hat{\lambda} = \lambda \cos(\hat{\phi}). \quad (9)$$

One of the problems that we will deal with is that we will not know with certainty which eigenvalue or eigenstate an estimate  $\hat{\lambda}$  really corresponds to. For example, an estimate  $\hat{\lambda}$  may correspond to an excited state, but the phase error can make it smaller than the ground state energy  $\lambda_0$ .

### 1. The quantum phase estimation circuit

A main building block of the quantum phase estimation algorithm we are going to use is the controlled walk operator

$$|0\rangle\langle 0| \otimes W^\dagger + |1\rangle\langle 1| \otimes W. \quad (10)$$

Compared to the controlled- $W$ , this operator needs the same number of gates to implement but doubles the resulting phase difference [32]. Using this version of the controlled walk operator, the controlled part in quantum phase estimation becomes

$$\sum_{k=0}^{N-1} |k\rangle\langle k| \otimes W^{2k-N}. \quad (11)$$

By using one more controlled  $W$  (rather than controlling between  $W$  and  $W^\dagger$ ) and relabelling, the controlled unitary becomes

$$\sum_{n=0}^{2N-1} |n\rangle\langle n| \otimes W^{n-N}. \quad (12)$$

Now suppose that the control register is initialized in the state  $|\Gamma\rangle = \sum_{n=0}^{2N-1} \gamma_n |n\rangle$  and the system register is initialized in a superposition of eigenstates  $|\Phi\rangle = \sum_j \Phi_j |\psi_j\rangle$ , then the state after applying the controlled unitary is

$$\sum_{n=0}^{2N-1} \gamma_n |n\rangle \otimes W^{n-N} |\Phi\rangle = \sum_j \Phi_j \sum_{n=0}^{2N-1} \gamma_n e^{i(n-N)\phi_j} |n\rangle \otimes |\psi_j\rangle. \quad (13)$$

After applying the QFT on the control register, the quantum state becomes

$$\sum_j \Phi_j \sum_{l=0}^{2N-1} (-1)^l e^{-(\phi_j - \pi l/N)/2} \Gamma\left(\phi_j - \frac{\pi l}{N}\right) |l\rangle \otimes |\psi_j\rangle, \quad (14)$$

where  $\Gamma(x)$  is a kernel function defined to be

$$\Gamma(x) = \frac{1}{\sqrt{2N}} \sum_{n=0}^{2N-1} e^{inx} \gamma_n. \quad (15)$$

The phase factor does not affect the probabilities so can be ignored in the following discussion.

### 2. The phase error

From Eq. (14) the probability of getting a phase estimate  $\pi l/N$  is

$$\Pr\left[\hat{\phi} = \frac{\pi l}{N}\right] = \sum_j |\Phi_j|^2 \left| \Gamma\left(\phi_j - \frac{\pi l}{N}\right) \right|^2. \quad (16)$$



Here  $\hat{\phi}$  is the random variable corresponding to the phase estimate. Note that the above probability resembles a convolution between two probability measures. From this observation we can write down a decomposition for  $\hat{\phi}$  in the following way:

$$\hat{\phi} = \phi + \Delta\phi, \quad (17)$$

where  $\phi$  is a random variable satisfying

$$\Pr[\phi = \phi_j] = |\Phi_j|^2, \quad (18)$$

which means that  $\phi$  is the output of an exact phase estimation for  $W$  and  $\Delta\phi$  is described by the following conditional distribution

$$\Pr\left[\Delta\phi = \frac{\pi l}{N} - \phi \middle| \phi\right] = \left|\Gamma\left(\phi - \frac{\pi l}{N}\right)\right|^2. \quad (19)$$

We call  $\Delta\phi$  the *phase error*.

In QPE, we want the phase error  $\Delta\phi$  to be small. One way to characterize this is through the variance of  $\Delta\phi$ , which is minimised by using amplitudes proportional to a cosine function [33]. The use of this control state for this application was considered in Ref. [32]. But oftentimes we want to make  $\Delta\phi$  small with a probability that is arbitrarily close to 1, thus giving us reliable estimates in many runs. More precisely, we want

$$\Pr[|\Delta\phi| \geq \epsilon_\phi | \phi] \leq \delta, \quad (20)$$

where  $|\Delta\phi|$  is calculated modulo  $2\pi$ . The task of optimising the performance is described by the theory of window functions, where the optimal performance is provided by the Slepian prolate spheroidal window [22, 23]. That is difficult to calculate, but can be approximated by the Kaiser window, which can be calculated by Bessel functions [25, 34]. Either of these can be used to obtain scaling  $\mathcal{O}(\epsilon_\phi^{-1} \ln(\delta^{-1}))$  with a small constant factor.

### 3. Window functions

In the theory of window functions, we would replace  $\gamma_n$  with a continuous function  $w(z)$ , so  $w(n - N + 1/2) = \gamma_n$ . Then the kernel function is approximately

$$\Gamma(x) = \frac{1}{\sqrt{2N}} \sum_{n=0}^{2N-1} e^{inx} \gamma_n \approx \frac{e^{i(N-1/2)x}}{\sqrt{2N}} \int_{-N}^N e^{izx} w(z) dz. \quad (21)$$

In the phase estimation it is also trivial to adjust the measurement so that a continuous range of outcomes is obtained. That is achieved by imposing an extra (known) phase shift in addition to  $\phi$ , and correcting for it in the estimate. Then we can consider the error probability distribution as a continuous function of  $x$  given by  $|\Gamma(x)|^2$ . When  $\Gamma(x)$  is approximated by the integral rather than the sum, then it is no longer periodic modulo  $\pi$ . The integral gives a non-periodic function of  $x$  over the whole real line.

The discrete case with  $\gamma_n$  corresponds to sampling  $w(z)$  at integer spacing at  $2N$  points from  $-(N - 1/2)$  to  $N - 1/2$ . That corresponds to multiplying  $w(z)$  by a comb function, so  $\Gamma(x)$  is the periodic function obtained by convolving the Fourier transform of  $w(z)$  with a comb function. As described in Ref. [25], the tail probabilities for this continuous case correspond to the average over the tail probabilities for the discrete case, where the samples are shifted (so starting from  $-(N - \nu)$  for  $\nu \in [0, 1]$ ). This means that some discrete case must give at least the performance (in terms of small tail probabilities) as the continuous case. A further advantage of using continuous window functions is that they can be scaled to a unit interval and analysed independently of the specific value of  $N$ .

Examples of the probability distributions for the error for a range of window functions are given in Fig. 2. The traditional textbook version [35] of quantum phase estimation uses a flat distribution, which corresponds to a control register of unentangled qubits in  $|+\rangle$  states. That window gives a narrow peak for the error that decays slowly, resulting in both large variance and large tails. The cosine window gives tails that decay more rapidly to yield excellent performance for the variance. The Kaiser and Slepian windows give tails that are lower resulting in smaller tail probabilities for confidence intervals. They do not decay as fast as the sine window, so the variance is larger. Note that for this example the Kaiser and Slepian windows are almost indistinguishable.



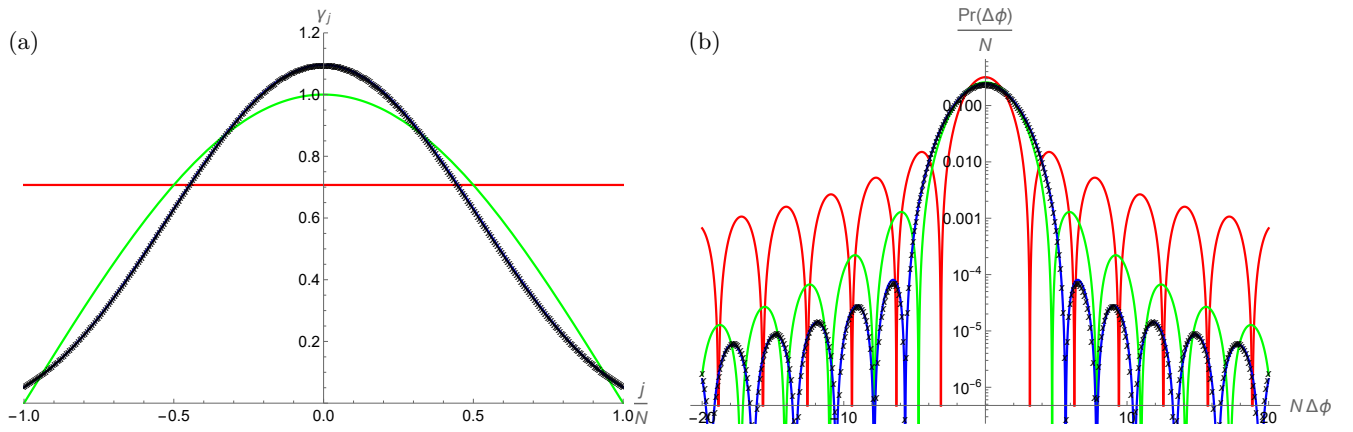


FIG. 2. The windows (a) and error probability distribution (b) for phase measurements. In red is the flat window, and in green is the cosine window, which provides the minimum phase variance. In blue is the Kaiser window with  $\alpha \approx 1.5$ , and the black crosses are the Slepian window with  $c = 5$ .

### III. MPS PREPARATION

In this section we analyse the Toffoli complexity needed for preparation of matrix product states. We first introduce a general unitary synthesis method that improves on the approach of LKS, then provide a method to generalise this approach to synthesising columns of a unitary, as is appropriate for MPS preparation.

#### A. New unitary synthesis method

Now we provide an alternative method of unitary synthesis that significantly improves on the method of LKS. We consider the rectangular array of beam splitters for decomposing a multiport interferometer as in Ref. [36]. This was previously considered for unitary synthesis in the optical context by Ref. [37], but we provide a significant improvement over that work.

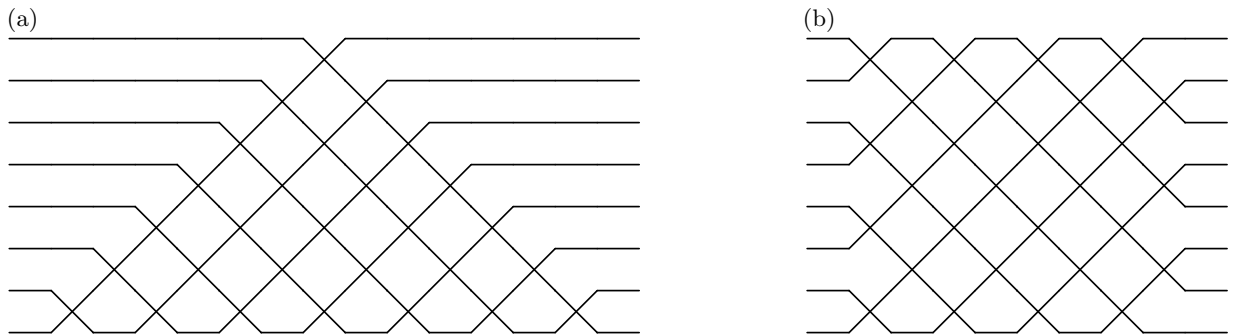


FIG. 3. The decomposition of a multiport interferometer for 8 modes, equivalent to a transformation on 3 qubits. The triangular Reck-Zeilinger [38] decomposition is in (a), and (b) is the rectangular decomposition from Ref. [36].

The form of the decomposition from Ref. [36] is shown in Fig. 3(b) for the example of  $N_{\text{un}} = 8$ . An  $M$ -port interferometer is equivalent to an  $N_{\text{un}} \times N_{\text{un}}$  unitary, and each layer of beam splitters is equivalent to a block-diagonal matrix with  $2 \times 2$  blocks. The first layer of beam splitters corresponds to a block-diagonal matrix where the first block is in rows and columns 1 and 2. The first layer of beam splitters in Fig. 3(b) would correspond to a unitary

matrix with nonzero entries shown by the asterisks in

$$\begin{bmatrix} * & * & & & & & \\ * & * & & & & & \\ & & * & * & & & \\ & & * & * & & & \\ & & & * & * & & \\ & & & * & * & & \\ & & & & * & * & \\ & & & & & * & * \\ & & & & & * & * \end{bmatrix}. \quad (22)$$

The next layer has blocks shifted by one, so would correspond to a matrix with nonzero entries

$$\begin{bmatrix} * & & & & & & \\ & * & * & & & & \\ & * & * & & & & \\ & & & * & * & & \\ & & & * & * & & \\ & & & & * & * & \\ & & & & * & * & \\ & & & & & * & \end{bmatrix}. \quad (23)$$

The layers alternate between operators of these two forms. Because the second form in Eq. (23) is equivalent to that in Eq. (22) except shifted by 1, it can be transformed to the same form by increment and decrement operations.

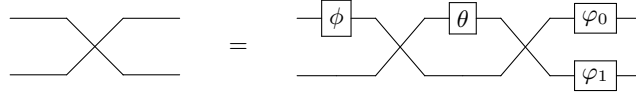


FIG. 4. The crossings depicted in Fig. 3 are general beam splitters with arbitrary phases and reflectivities (left). They can be implemented with two 50/50 beam splitters with phase shifts (right).

In the optical interferometer, each beam splitter may be expressed as two 50/50 beam splitters with a phase shift in between, as in Fig. 4. The 50/50 beam splitter corresponds to a Hadamard matrix. That is equivalent to expressing a general qubit unitary as

$$\begin{bmatrix} e^{i\varphi_0} & 0 \\ 0 & e^{i\varphi_1} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} e^{i\theta} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{bmatrix}. \quad (24)$$

Since the layers correspond to  $2 \times 2$  block-diagonal unitaries, this decomposition may be performed on the blocks individually, so for example Eq. (22) is of the form (with  $H$  the Hadamard)

$$\begin{bmatrix} e^{i\varphi_0} & & & & & & & \\ & e^{i\varphi_1} & & & & & & \\ & & e^{i\varphi_2} & & & & & \\ & & & e^{i\varphi_3} & & & & \\ & & & & e^{i\varphi_4} & & & \\ & & & & & e^{i\varphi_5} & & \\ & & & & & & e^{i\varphi_6} & \\ & & & & & & & e^{i\varphi_7} \end{bmatrix} \begin{bmatrix} H & & & \\ & H & & \\ & & H & \\ & & & H \end{bmatrix} \begin{bmatrix} e^{i\theta_0} & & & & & & \\ & 1 & & & & & \\ & & e^{i\theta_1} & & & & \\ & & & 1 & & & \\ & & & & e^{i\theta_2} & & \\ & & & & & 1 & \\ & & & & & & e^{i\theta_3} \\ & & & & & & & 1 \end{bmatrix} \begin{bmatrix} H & & & \\ & H & & \\ & & H & \\ & & & H \end{bmatrix} \begin{bmatrix} e^{i\phi_0} & & & & & & \\ & 1 & & & & & \\ & & e^{i\phi_1} & & & & \\ & & & 1 & & & \\ & & & & e^{i\phi_2} & & \\ & & & & & 1 & \\ & & & & & & e^{i\phi_3} \\ & & & & & & & 1 \end{bmatrix}. \quad (25)$$

This can be further simplified, because the phases  $\varphi_j$  can be combined with the next layer. Then the next layer can be similarly decomposed, and its  $\varphi_j$  phases can be combined with the layer after, and so on. As a result, the phases on *all* basis states are only needed for the very last operation. As a result we have  $2N_{\text{un}}$  layers of phase shifts with  $\lfloor N_{\text{un}}/2 \rfloor$  phases each, and one final layer with  $N_{\text{un}}$  phases. Then the block diagonal matrices have diagonals with all Hadamards, which can be achieved with just a Hadamard on a single qubit.

Note that the Toffoli cost of QROM is minimised if we are able to output more of the data together. It is possible to output the data for two layers of phase shifts at once, so there are  $N_{\text{un}} + 1$  uses of QROM with output size  $N_{\text{un}}$ . Considering the first layer of beam splitters, it is now decomposed into two layers of phase shifts that can be chosen to be only on odd (or even) modes, as well as two layers of 50/50 beam splitters. The equivalent quantum circuit corresponds to two layers of phase shifts that only depend on the first  $n - 1$  qubits, and two layers of Hadamards on the last qubit, as shown in Fig. 5.

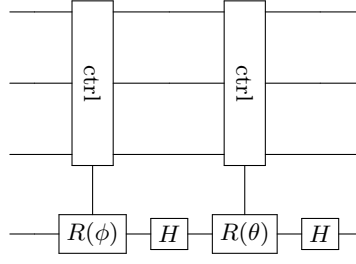


FIG. 5. Two layers of phases controlled by the first  $n - 1$  qubits and Hadamards on the last qubit. The boxes labelled  $R(\phi)$  and  $R(\theta)$  are phase rotations controlled by the first  $n - 1$  qubits.

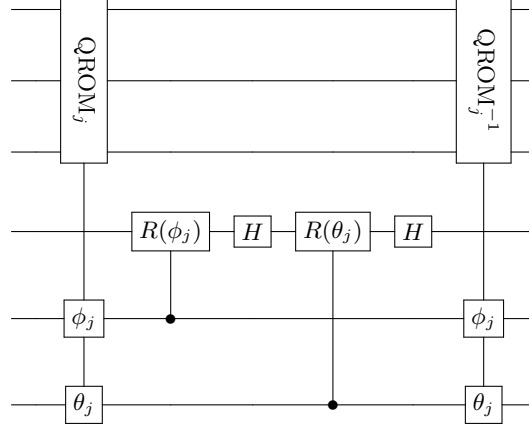


FIG. 6. Two layers of phases obtained by using a single QROM. The bottom two registers are temporary data registers used for the output of the QROM. The QROM on the left outputs both  $\phi_j$  and  $\theta_j$ , then the QROM on the right is an inverse QROM for erasure. In the middle the rotations are controlled by the values in the data registers.

It is therefore possible to perform a QROM on the first  $n - 1$  qubits, perform the operations on qubit  $n$ , then erase the QROM because it does not depend on the last qubit. The quantum circuit is shown in Fig. 6. A further simplification is possible because of the way QROM erasure is performed. The data qubits are measured in the  $X$  basis, and sign corrections need to be performed on the control qubits. However, these sign corrections are phases that can be combined into  $\varphi_j$ . That is, we just need to modify the phases used in the remainder of the circuit to account for the sign fixups needed for the QROM erasure. The Toffoli complexity is therefore

$$\left\lceil \frac{N_{\text{un}}}{2\Lambda} \right\rceil + (\Lambda - 1)2b - 1 \quad (26)$$

to output the  $N_{\text{un}}/2$  items of data of size  $2b$  (for two rotations), with  $\Lambda$  a power of 2. In this expression  $-1$  accounts for the fact that the cost in the first term is for unary iteration [39], and the Toffoli cost of unary iteration is 1 less than the number of items [32]. This is a cost including a control, and this control will be needed in the overall scheme so is allowed for here. The cost for the two rotations is  $2(b - 2)$ , where the  $b - 2$  cost for a phase rotation is explained in Ref. [40]. In the following we will bundle these two costs together.

A minor issue is the even layers where the blocks are shifted by 1. This is easily accounted for in the quantum circuit by applying an increment in the computational basis to shift the blocks such that the layer may be implemented in the same way as for the first layer. Each increment (or decrement) can be performed with  $n - 2$  Toffolis, and there are  $N_{\text{un}} - 1$  needed, for a total complexity of  $(n - 2)(N_{\text{un}} - 1)$ . (Recall that the cost of modular addition is  $n - 1$  as per Ref. [41], with a saving of one Toffoli when the number to be added is classical, rather than provided in a quantum register.) Bringing all these complexities together we have the following.

- The complexity for the  $2N_{\text{un}}$  layers of phase shifts and Hadamards is

$$N_{\text{un}} \left( \left\lceil \frac{N_{\text{un}}}{2\Lambda} \right\rceil + 2\Lambda b - 5 \right). \quad (27)$$

- The increments and decrements have complexity  $(n - 2)(N_{\text{un}} - 1)$ .

- The final phase shifts have complexity

$$\left\lceil \frac{N_{\text{un}}}{\Lambda} \right\rceil + \Lambda b + \left\lceil \frac{N_{\text{un}}}{\Lambda'} \right\rceil + \Lambda' - 6. \quad (28)$$

For Eq. (28) there is a term  $-6$  which comes from  $-1$  for the initial QROM,  $-2$  for the addition into the phase gradient, and  $-3$  for the sign fixup. The sign fixup is explained in Fig. 6 of Ref. [39], and can be constructed from two unary iterations on subsets of the qubits. Only one of these need be controlled, for an overall saving of  $-3$  Toffolis. In numerical testing of this approach against the LKS approach, we find that the complexity is reduced by about a factor of 7 over a wide range of parameters.

### B. Synthesis of columns of unitary

We can now use this approach to reduce the cost for the case where it is only necessary to synthesise half the columns of the unitary. The key to this simplification is that the unitary can be simplified by an initial unitary on dimension  $N_{\text{un}}/2$  (ignoring the extra qubit), then afterward using the extra qubit to control one of two unitaries on the remaining qubits. This then diagonalises the two  $N_{\text{un}}/2 \times N_{\text{un}}/2$  blocks of the unitary that we need. That is, we have a decomposition of the unitary as

$$\begin{bmatrix} A & ? \\ B & ? \end{bmatrix} = \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} D_1 & ? \\ D_2 & ? \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & V \end{bmatrix}. \quad (29)$$

In this expression the question marks indicate blocks that we do not need to specify, so  $A$  and  $B$  indicate half of the columns of the unitary matrix that need to be correctly produced. The blocks  $D_1, D_2$  are diagonal, and  $U_1, U_2, V$  are unitaries. That is, we require

$$A = U_1 D_1 V, \quad B = U_2 D_2 V. \quad (30)$$

The matrices  $U_1, D_1, V$  are easily determined via a singular value decomposition of  $A$ . Then  $U_2, D_2$  can be determined from a QR decomposition of  $BV^\dagger$ . The QR decomposition guarantees that  $D_2$  is upper triangular, and the requirement that  $D_1$  and  $D_2$  are blocks of a unitary matrix ensures that  $D_2$  is diagonal.

So, the procedure to apply the unitary that we need is as follows.

1. Perform a unitary of dimension  $N_{\text{un}}/2 \times N_{\text{un}}/2$  on the first  $n - 1$  qubits.
2. Use  $n - 1$  qubits to control rotation on the remaining qubit.
3. Use that qubit to control unitaries on the  $n - 1$  qubits.

For the costing of this procedure, the cost of the first step is the same as in the formulae for unitary synthesis above, except replacing  $N_{\text{un}}$  with  $N_{\text{un}}/2$  and  $n$  with  $n - 1$ .

However, in the MPS preparation we have a unitary on these  $n - 1$  qubits first, and this unitary can be combined with that one so it need not be performed and requires no additional complexity. This principle is illustrated in Fig. 7, where it can be seen that the dimension  $N_{\text{un}}/2$  register where  $U_1$  and  $U_2$  are performed is the same as that where  $V'$  is applied for the next step.

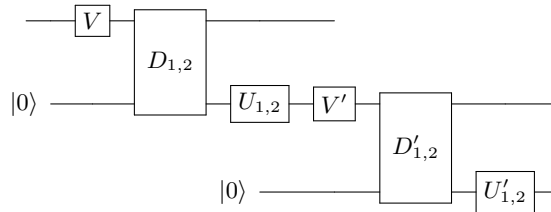


FIG. 7. Two consecutive steps in the MPS state preparation, showing how the unitaries required for consecutive steps can be combined. The primes show the unitaries needed for the second step. The input registers with  $|0\rangle$  are qubits, and the top two outputs are qubits.

Step 2 is just a QROM and rotation, for complexity

$$\left\lceil \frac{N_{\text{un}}}{2\Lambda} \right\rceil + \Lambda b - 2. \quad (31)$$

The  $-2$  term is because the rotation is implemented by a controlled addition or subtraction on half the phase, which requires one more bit for the addition into the phase gradient register. The QROM erasure does not introduce a Toffoli cost here, because the sign fixup can be incorporated into the unitaries  $U_1$  and  $U_2$ . A subtlety in this costing is that the blocks are distinguished by the *least* significant qubit, so we only need consider dimension  $N_{\text{un}}/2$  for the sub-blocks. For simplicity we are assuming  $N_{\text{un}}$  is even for this discussion.

For the controlled unitaries on dimension  $N_{\text{un}}/2$ , in each layer we have the same size of the QROM as for the operations on the full dimension  $N_{\text{un}}$ , with the only difference being that we now only need  $N_{\text{un}}/2$  layers rather than  $N_{\text{un}}$ . Also, the increments and decrements are on  $n-1$  qubits rather than  $n$ , slightly decreasing the complexity. The final phase shifts also have the same cost as before. The complete cost is therefore

$$\left\lceil \frac{N_{\text{un}}}{2\Lambda} \right\rceil + \Lambda b - 2 \quad (\text{controlled qubit rotation}) \quad (32)$$

$$+ \frac{N_{\text{un}}}{2} \left( \left\lceil \frac{N_{\text{un}}}{2\Lambda} \right\rceil + 2\Lambda b - 5 \right) \quad (\text{phasing layers}) \quad (33)$$

$$+ (n-3)(N_{\text{un}}/2 - 1) \quad (\text{increments and decrements}) \quad (34)$$

$$+ \left\lceil \frac{N_{\text{un}}}{\Lambda} \right\rceil + \Lambda b + \left\lceil \frac{N_{\text{un}}}{\Lambda'} \right\rceil + \Lambda' - 6. \quad (\text{final phase shifts}) \quad (35)$$

This gives a similar factor of  $\sim 7$  improvement over LKS as for synthesis of the complete unitary.

So far we have discussed the case where the dimension on each site is  $d=2$  for the MPS. We can solve the case for  $d>2$  by iterating the procedure, and the cost is multiplied by  $d-1$ . To explain the principle, we first explain how to obtain the result for  $d=3$ . Then we can construct a decomposition of the form

$$\begin{bmatrix} A & ? & ? \\ B & ? & ? \\ C & ? & ? \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & U_{11} & U_{12} \\ 0 & U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} A & ? & ? \\ R_{11} & R_{12} & ? \\ 0 & R_{22} & ? \end{bmatrix}. \quad (36)$$

In this expression,  $A, B, C$  are the blocks of the unitary we wish to construct,  $U_{ij}$  are blocks of a unitary, and  $R_{ij}$  are blocks of an upper triangular matrix. The block  $R_{21}$  must be zero because of the upper triangular form, so is just given as zero above. In this form,  $U_{ij}$  and  $R_{ij}$  are obtained by a QR decomposition of the 4 blocks of the desired unitary operation including  $B, C$  and two unspecified blocks indicated by the question marks. It doesn't matter that the blocks indicated by question marks are unknown, because they can just be replaced with zero for the purposes of the QR decomposition. We only need to determine  $R_{11}$  and  $U_{11}, U_{21}$ , because the other blocks do not affect  $A, B, C$ , and those blocks are obtained correctly by the QR decomposition with question marks replaced with zero.

Now it is possible to correctly apply the blocks  $U_{11}$  and  $U_{21}$  of the first unitary by the above procedure for synthesising half of the columns of a unitary. Similarly,  $A$  and  $R_{11}$  have orthonormal columns, so can be regarded as half the columns of some unitary, which can also be synthesised by the above scheme. Moreover, the construction of the unitary with blocks  $U_{11}, U_{21}$  requires an initial unitary of dimension  $N_{\text{un}}/3$ . This unitary can be combined with  $R_{11}$  and so need not add to the cost of synthesising  $U_{11}, U_{21}$ . Similarly, if this unitary synthesis is part of MPS preparation, the initial dimension  $N_{\text{un}}/3$  unitary for synthesis of  $A, R_{11}$  can be combined with other operations and does not add to the cost.

In the general case, we need to construct the correct first  $\chi$  columns of a matrix of size  $N_{\text{un}} = d\chi$ . Consider the block consisting of the last  $(d-1)\chi$  rows and first  $(d-1)\chi$  columns (equivalent to  $B, C$  and the question marks in the  $d=3$  example above). Apply a QR decomposition, to express it in the form of a unitary operation followed by an upper triangular matrix. Similar to the example above, the question marks may be replaced with zero in this decomposition, because they only affect blocks that we do not need to specify.

In exactly the same way as in the example above, we need only correctly synthesise the first  $\chi$  columns of the first operation, and the upper-triangular form guarantees that the only non-zero blocks are  $A$  and an upper triangular  $\chi \times \chi$  block (which is  $R_{11}$  in the example above). These blocks may be synthesised by the above procedure for  $d=2$ , where we are synthesising half the columns of a unitary.

The unitary which we obtained by the QR decomposition is of size  $(d-1)\chi \times (d-1)\chi$ , and we only need to correctly reproduce the first  $\chi$  columns. In the example above, these were  $U_{11}, U_{21}$ . Therefore we have reduced the problem to synthesising the first  $\chi$  columns of a  $(d-1)\chi \times (d-1)\chi$  unitary, which is the same as the initial problem, with  $d$  reduced by 1. We may therefore iterate this procedure to completely reduce the problem to that for  $d=2$ . Therefore, we are able to reduce the problem to  $d-1$  applications of the scheme for  $d=2$ .

There is a very small increase in the cost for  $d>4$ . Because the schemes for  $d=2$  are on a subspace, the QROMs need to be controlled, which is a cost accounted for above. However, for  $d>4$  we also need Toffolis to produce the

qubit flagging the control for the QROMs. That is only performed once for each application of the  $d = 2$  scheme, so is a negligible contribution to the overall cost.

We can also combine the  $d = 2$  schemes in a slightly more efficient way. First, note that at each step there is a controlled unitary performed on two dimension  $\chi$  subspaces, but only one is used in the next  $d = 2$  scheme. Instead of performing both, we can instead just perform one unitary of dimension  $\chi \times \chi$ , and then perform controlled unitaries on the  $d$  subspaces at the end. Then the costs of  $d - 1$  controlled  $U_1, U_2$  operations are replaced with  $d - 2$  unitaries of dimension  $\chi \times \chi$ , followed by selection between  $d$  unitaries of this size.

In either case the increase in cost with  $d$  is linear in  $d$ , whereas for the LKS approach the factor is about  $\sqrt{d/2}$ , so LKS has better scaling with  $d$ . For  $d = 4$  we find that the improvement over LKS is about a factor of 3.5, and the crossover where LKS is more efficient is for large values of  $d$  about 30.

It is also possible to prepare states more efficiently than the procedure of LKS. The method is to use interspersed layers of Hadamards and diagonal phasing operators, and we find that three phasing layers are sufficient. That is, prepare an equal superposition state, apply phases in the computational basis, a layer of Hadamards, then more phases. The initial two layers of phases and Hadamards produce the correct amplitudes. The phases may be found efficiently by a simple generalisation of the Gerchberg-Saxton algorithm; we have tested this with dimension up to  $2^{20}$ . This approach may also be used in the MPS preparation to slightly reduce the cost. Increasing the number of layers can also be used for unitary synthesis, but solving for the phases becomes computationally intractable for larger dimensions (above about 128).

#### IV. PHASE ESTIMATION WITH WINDOW FUNCTIONS

First we describe how to perform phase estimation that is optimal for confidence intervals. That is, for a given number of controlled applications of an operator and confidence level, it gives the smallest width of the confidence interval. This problem is related to that of window functions in classical signal processing theory, and optimal confidence intervals are given by the prolate spheroidal window. The analysis of the prolate spheroidal window was given in Ref. [23], but it is difficult to calculate, so Ref. [25] gives an analysis of the Kaiser window, which gives near optimal results. The use of the Kaiser window was also mentioned on page 35 of Ref. [40]. The error for the Kaiser window is in terms of the sinc function, so can be calculated with standard mathematical software. The prolate spheroidal window can also be calculated using specialist mathematical software.

We will provide both asymptotic results for the cost of phase estimation using these window functions, as well as exact expressions using special functions and numerical results calculated from these functions. For the Kaiser window, we provide first-order approximations for the error and cost in Eq. (41) and Eq. (44), respectively, and higher-order approximations for the error in Eq. (50), and cost in Eq. (52). For the prolate spheroidal window, we provide a higher-order approximation of the error in Eq. (59), and the cost in Eq. (60). We show that, although the Slepian window provides asymptotically improved error, the cost is the same up to leading order.

##### A. The Kaiser window

First, we summarise the Kaiser window and its asymptotic scaling. The standard form of the Kaiser window is proportional to

$$w(x) = I_0 \left( \pi \alpha \sqrt{1 - (x/N)^2} \right), \quad (37)$$

for  $|x| \leq N$ , and 0 otherwise. As discussed in Section II B 3, the control state used would correspond to samples of this continuous window at discrete points. The window function yields a probability distribution for the error  $\theta$  proportional to

$$\frac{\sin^2 \left( \sqrt{(N\theta)^2 - (\pi\alpha)^2} \right)}{(N\theta)^2 - (\pi\alpha)^2}. \quad (38)$$

That is the square of the Fourier transform of  $w(x)$ .

A simple approximation for the tail probabilities with the Kaiser window is given in Ref. [25]. The method used there is to first approximate the normalisation by approximating the centre of the distribution by a Gaussian. In that approximation, the integral over  $\theta$  then gives

$$\int_{-\infty}^{\infty} \frac{\sinh^2(\pi\alpha)}{\pi^2 \alpha^2} e^{-N^2(\pi\alpha \coth(\pi\alpha) - 1)\theta^2/(\pi^2 \alpha^2)} d\theta = \frac{\sinh^2(\pi\alpha)}{N\sqrt{\pi\alpha}\sqrt{\pi\alpha \coth(\pi\alpha) - 1}}$$

$$\begin{aligned}
&\approx \frac{e^{2\pi\alpha}}{4N\sqrt{\pi\alpha}\sqrt{\pi\alpha-1}} \\
&\approx \frac{e^{2\pi\alpha}}{4N\pi\alpha^{3/2}}, \tag{39}
\end{aligned}$$

where in the second line we have used  $\sinh^2(\pi\alpha) \approx e^{2\pi\alpha}/4$  and  $\coth(\pi\alpha) \approx 1$ , and in the last line we have made an approximation for large  $\alpha$ .

Then the tail probabilities can be approximated by replacing  $\sin^2$  with  $1/2$ , and integrating from the first zero at  $\theta = (\pi/N)\sqrt{1+\alpha^2}$  to give

$$\int_{(\pi/N)\sqrt{1+\alpha^2}}^{\infty} \frac{1}{[N^2\theta^2 - (\pi\alpha)^2]} d\theta = \frac{\operatorname{arcsinh}(\alpha)}{\pi N\alpha} \approx \frac{\ln(2\alpha)}{\pi N\alpha}, \tag{40}$$

using  $\operatorname{arcsinh}(\alpha) \approx \ln(2\alpha)$ . Dividing by the approximation for the normalisation in Eq. (39) then gives

$$\delta \approx 4 \ln(2\alpha) \sqrt{\alpha} e^{-2\pi\alpha}. \tag{41}$$

This is a factor of 2 smaller than the expression in Ref. [25], because we have approximated  $\sin^2$  with  $1/2$ .

Therefore, to obtain confidence level  $1 - \delta$  (so the probability of error outside the range is  $\delta$ ), we take

$$\ln(1/\delta) \approx 2\pi\alpha - \ln[4 \ln(2\alpha) \sqrt{\alpha}]. \tag{42}$$

Solving for  $\alpha$  then gives the approximation

$$\alpha \approx \frac{1}{2\pi} \ln(1/\delta) + \frac{1}{4\pi} \ln(8 \ln(4/\delta)/\pi) + \frac{1}{2\pi} \ln[\ln(\ln(4/\delta)/\pi)]. \tag{43}$$

The size of the confidence interval is  $(\pi/N)\sqrt{1+\alpha^2}$ , so if that needs to be  $\epsilon$ , we should take

$$N = \frac{\pi}{\epsilon} \sqrt{1+\alpha^2} = \frac{1}{2\epsilon} \ln(1/\delta) + \mathcal{O}(\epsilon^{-1} \ln \ln(1/\delta)). \tag{44}$$

The higher-order  $\ln \ln$  term for  $\alpha$  is larger than the correction term for approximating  $\sqrt{1+\alpha^2}$  by  $\alpha$ .

## B. Higher-order terms

It is possible to obtain more accurate approximations by taking advantage of special functions. First the integral can be given as, using the Plancherel theorem,

$$\int \frac{\sin^2 \left( \sqrt{N^2\theta^2 - (\pi\alpha)^2} \right)}{N^2\theta^2 - (\pi\alpha)^2} d\theta = \frac{\pi}{2N} \int_{-1}^1 I_0^2 \left( \pi\alpha \sqrt{1-x^2} \right) dx. \tag{45}$$

Ignoring the factor of  $1/N$  for simplicity, this integral can be evaluated as

$$\frac{\pi}{2} \int_{-1}^1 I_0^2 \left( \pi\alpha \sqrt{1-x^2} \right) dx = \pi \int_0^1 I_0^2(\pi\alpha y) \frac{y dy}{\sqrt{1-y^2}} = \frac{\pi}{2} [I_0(2\pi\alpha)(2 + \pi L_1(2\pi\alpha)) - \pi I_1(2\pi\alpha)L_0(2\pi\alpha)], \tag{46}$$

where  $L_0, L_1$  are modified Struve functions. Expanding about  $\alpha = \infty$  then gives

$$\frac{e^{2\pi\alpha}}{4\pi\alpha^{3/2}} \left( 1 + \frac{5}{2^4\pi\alpha} + \frac{129}{2^9\pi^2\alpha^2} + \frac{2655}{2^{13}\pi^3\alpha^3} + \frac{301035}{2^{19}\pi^4\alpha^4} + \frac{10896795}{2^{23}\pi^5\alpha^5} + \frac{961319205}{2^{29}\pi^6\alpha^6} + \mathcal{O}(\alpha^{-7}) \right). \tag{47}$$

For the integral over the tails, we can use

$$\begin{aligned}
2 \int_{(\pi/N)\sqrt{1+\alpha^2}}^{\infty} \frac{\sin^2 \sqrt{N^2\theta^2 - \pi^2\alpha^2}}{N^2\theta^2 - \pi^2\alpha^2} d\theta &= \frac{2}{N} \int_{\pi\sqrt{1+\alpha^2}}^{\infty} \frac{\sin^2 \sqrt{\theta^2 - \pi^2\alpha^2}}{\theta^2 - \pi^2\alpha^2} d\theta \\
&= \frac{2}{N} \int_{\pi}^{\infty} \frac{\sin^2 x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx
\end{aligned}$$



$$\begin{aligned}
&= \frac{1}{N} \int_{\pi}^{\infty} \frac{1 - \cos 2x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx \\
&= \frac{\operatorname{arcsinh}(\alpha)}{\pi N \alpha} - \frac{1}{N} \int_{\pi}^{\infty} \frac{\cos 2x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx.
\end{aligned} \tag{48}$$

For the remaining integral, we can use integration by parts to give, where Ci is the cosine integral,

$$\begin{aligned}
- \int_{\pi}^{\infty} \frac{\cos 2x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx &= \frac{\operatorname{Ci}(2\pi)}{\pi\sqrt{1 + \alpha^2}} - \int_{\pi}^{\infty} \frac{x \operatorname{Ci}(2x)}{(x^2 + \pi^2\alpha^2)^{3/2}} dx \\
&= \frac{\operatorname{Ci}(2\pi)}{\pi\sqrt{1 + \alpha^2}} + \frac{4\pi^2 \operatorname{Ci}(2\pi) - 1}{8\pi^3(1 + \alpha^2)^{3/2}} + \int_{\pi}^{\infty} \frac{3x[\cos(2x) + 2x \sin(2x) - 4x^2 \operatorname{Ci}(2x)]}{8(x^2 + \pi^2\alpha^2)^{5/2}} dx.
\end{aligned} \tag{49}$$

Note that the final integral here is of order  $\alpha^{-5}$ , and repeating the integration by parts makes the remaining integral higher and higher order in  $\alpha$ . We give further details of the integration by parts in Appendix A. Dividing by the asymptotic expression above for the normalisation factor in Eq. (47) gives, with  $C_{\alpha} = \ln(2\alpha) + \operatorname{Ci}(2\pi)$ ,

$$\delta = 4C_{\alpha}\sqrt{\alpha}e^{-2\pi\alpha} \left[ 1 - \frac{5}{16\pi\alpha} + \mathcal{O}(\alpha^{-2}) \right]. \tag{50}$$

The leading-order term here only differs from what we obtained in Eq. (41) in that it has  $C_{\alpha}$  with  $\operatorname{Ci}(2\pi)$  in addition to  $\ln(2\alpha)$ .

Given error  $\delta$ , expanding in a series solution for  $\alpha$  yields

$$\alpha = \frac{\ln(1/\delta)}{2\pi} + \frac{\ln(8 \ln(4/\delta)/\pi)}{4\pi} + \frac{1}{2\pi} \ln[\operatorname{Ci}(2\pi) + \ln(\ln(4/\delta)/\pi)] + \mathcal{O}\left(\frac{\ln \ln(1/\delta)}{\ln(1/\delta)}\right). \tag{51}$$

The leading-order terms here only differ from Eq. (43) in that there is an extra  $\operatorname{Ci}(2\pi)$  in the third term, so the previous approximation given can be expected to be accurate. The expression for  $N$  is then

$$N = \frac{\pi}{\epsilon} \sqrt{1 + \alpha^2} = \frac{\ln(1/\delta)}{2\epsilon} + \frac{\ln(8 \ln(4/\delta)/\pi)}{4\epsilon} + \frac{1}{2\epsilon} \ln[\operatorname{Ci}(2\pi) + \ln(\ln(4/\delta)/\pi)] + \mathcal{O}\left(\frac{\ln \ln(1/\delta)}{\epsilon \ln(1/\delta)}\right). \tag{52}$$

To this order of approximation  $\sqrt{1 + \alpha^2} \approx \alpha$ , and the correction is in the order term above.

The performance of the Kaiser window can be improved by adjusting the width from  $\pi\sqrt{1 + \alpha^2}$  to  $\pi\sqrt{\Delta^2 + \alpha^2}$  for some general  $\Delta \neq 1$ . Then the integral over the tails can be adjusted to

$$2 \int_{(\pi/N)\sqrt{\Delta^2 + \alpha^2}}^{\infty} \frac{\sin^2 \sqrt{N^2\phi^2 - \pi^2\alpha^2}}{(N^2\phi^2 - \pi^2\alpha^2)} d\phi = \frac{\operatorname{arcsinh}(\alpha/\Delta)}{\pi N \alpha} - \frac{1}{N} \int_{\pi\Delta}^{\infty} \frac{\cos 2x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx. \tag{53}$$

For the remaining integral,

$$\begin{aligned}
&- \int_{\pi\Delta}^{\infty} \frac{\cos 2x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx \\
&= \frac{\operatorname{Ci}(2\pi\Delta)}{\pi\sqrt{\Delta^2 + \alpha^2}} - \int_{\pi\Delta}^{\infty} \frac{x \operatorname{Ci}(2x)}{(x^2 + \pi^2\alpha^2)^{3/2}} dx \\
&= \frac{\operatorname{Ci}(2\pi\Delta)}{\pi\sqrt{\Delta^2 + \alpha^2}} + \frac{4\pi^2\Delta^2 \operatorname{Ci}(2\pi) - \cos(2\pi\Delta) - 2\pi\Delta \sin(2\pi\Delta)}{8\pi^3(\Delta^2 + \alpha^2)^{3/2}} + \int_{\pi\Delta}^{\infty} \frac{3x[\cos(2x) + 2x \sin(2x) - 4x^2 \operatorname{Ci}(2x)]}{8(x^2 + \pi^2\alpha^2)^{5/2}} dx,
\end{aligned} \tag{54}$$

with further details given in Appendix A. Dividing by the normalisation factor in Eq. (47) and using  $C_{\alpha,\Delta} = \ln(2\alpha/\Delta) + \operatorname{Ci}(2\pi\Delta)$ , we obtain

$$\delta = 4C_{\alpha,\Delta}\sqrt{\alpha}e^{-2\pi\alpha} \left[ 1 - \frac{5}{16\pi\alpha} + \mathcal{O}(\alpha^{-2}) \right], \tag{55}$$

which is the same as in Eq. (50) except for the more general expression for  $C_{\alpha,\Delta}$  and the higher-order terms not shown here but given in Appendix A.

In practice we would consider a confidence interval of fixed width  $c = \pi\sqrt{\Delta^2 + \alpha^2}$  and aim to minimise the cost. If we expand in a series in  $c$ , then we obtain

$$\delta = 4C_{c,\Delta}\sqrt{c/\pi}e^{-2c} \left[ 1 - \frac{(5 - 16\pi^2\Delta^2)}{16c} + \mathcal{O}(c^{-2}) \right], \tag{56}$$

where  $C_{c,\Delta} = \ln(2c/\pi\Delta) + \text{Ci}(2\pi\Delta)$ . Now  $C_{c,\Delta}$  takes its minimum value at  $\Delta = 1$ , but the second term in square brackets above increases with  $\Delta$ . This implies that for  $c$  not too large, the optimal value of  $\Delta$  will be less than 1, but in the limit of large  $c$  the optimal value of  $\Delta$  approaches 1. That is indeed what is found numerically. This also implies that the first terms for  $N$  in terms of  $\epsilon$  in Eq. (52) are appropriate even when optimising  $\Delta$ . In fact, optimising  $\Delta$  would make the third term *larger*, with the improvement only in the higher-order terms. This is because the third term is equivalent to  $\ln(C_{c,\Delta})/2\epsilon$ . It is more accurate to continue using the third term as given, since the true value of  $N$  is reduced when optimising  $\Delta$ , and using a larger value of  $C_{c,\Delta}$  in this term would make the estimate of  $N$  larger.

### C. The prolate spheroidal window

For comparison with the optimal window, the error is given in Eq. (13) of [23] as

$$\delta = 4\sqrt{\pi c} e^{-2c} \left[ 1 - \frac{3}{32c} + \mathcal{O}(c^{-2}) \right]. \quad (57)$$

That result is from Eq. (1.36) of Ref. [22], with  $n = 0$ . In fact, Eq. (4.4) of that work gives

$$\delta = 4\sqrt{\pi c} e^{-2c} \left[ 1 - \frac{3}{32c} - \frac{389}{2^{11}c^2} + \mathcal{O}(c^{-3}) \right], \quad (58)$$

where we have used  $n = 0$  and simplified  $2334/3! = 389$  from the expression in that work. It turns out that this expression is not correct. By repeating the derivation as given on page 138 of Ref. [22], we obtain

$$\delta = 4\sqrt{\pi c} e^{-2c} \left[ 1 - \frac{7}{16c} - \frac{91}{2^9 c^2} + \mathcal{O}(c^{-3}) \right]. \quad (59)$$

See Appendix B for the details of this derivation.

In this expression the factor of  $\ln c$  that was obtained for the Kaiser window error has been eliminated, so this error has asymptotically better scaling. According to the above analysis for the asymptotic expansion for  $N$ , we would get

$$N = \frac{c}{\epsilon} = \frac{\ln(1/\delta)}{2\epsilon} + \frac{\ln(8\pi \ln(4\sqrt{\pi}/\delta))}{4\epsilon} + \mathcal{O}\left(\frac{\ln(\ln(1/\delta))}{\epsilon \ln(1/\delta)}\right). \quad (60)$$

That is, despite the optimal window giving asymptotically smaller error, the cost  $N$  is reduced by only removing the third term in Eq. (52). That term is triple-logarithmic in  $1/\delta$ , meaning that the performance is only marginally improved.

The optimal window is the angular spheroidal function of the first kind  $\text{PS}_{0,0}(c, z)$  for  $z \in [-1, 1]$ . Integrating then gives (see Appendix B for explanation of why this integral is used)

$$\begin{aligned} (1 - \delta)\text{PS}_{0,0}(c, 0) &= \frac{c}{\pi} \int_{-1}^1 \text{sinc}(cz) \text{PS}_{0,0}(c, z) dz \\ &= \frac{2c}{\pi} \text{PS}_{0,0}(c, 0) [S_{0,0}^1(c, 1)]^2, \end{aligned} \quad (61)$$

so

$$\delta = 1 - \frac{2c}{\pi} [S_{0,0}^1(c, 1)]^2, \quad (62)$$

where the function  $S_{0,0}^1(c, 1)$  is the radial spheroidal function of the first kind. Thus it is possible to compute the error in terms of special functions. Various methods are discussed in Appendix C.

In comparison, the value of  $N$  for root-mean-square (RMS) error  $\epsilon$  is [32]

$$N \approx \frac{\pi}{2\epsilon}. \quad (63)$$

That is, to leading order the expression for  $N$  for the confidence interval replaces  $\pi$  with  $\ln(1/\delta)$ . For a 95% confidence interval, for example,  $\ln(1/\delta)$  is less than  $\pi$ , but calculation using the exact expression in Eq. (62) shows that the complexity is about 63% larger than for achieving RMS error  $\epsilon$ . For a 90% confidence interval the complexity is only about 35% larger than for achieving RMS error  $\epsilon$ .

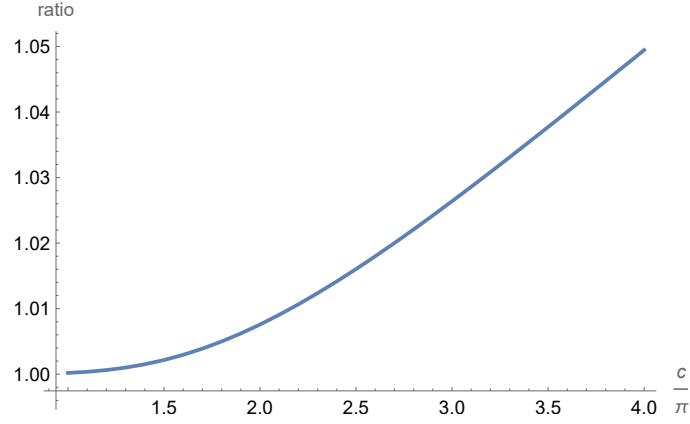


FIG. 8. The ratio of the error for the Kaiser window to that for the optimal Slepian window as a function of  $c$ . For the Kaiser window the value of  $\Delta$  is optimised to minimise the error.

Next we numerically compare the error outside the confidence interval for the Kaiser window versus that for the optimal window. According to the above asymptotic analysis for the error, the ratio of the errors should increase with  $c$ , and we find that occurs even optimising for  $\Delta$ . In Fig. 8 we show the ratio, and it is very close to 1 for  $c = \pi$ , and increases to be about 5% larger for  $c = 4\pi$ .

Now we test the series of Slepian compared to the one we have given in Eq. (59). In Fig. 9 we take the difference between the exact error and asymptotic approximation divided by  $4\sqrt{\pi c} e^{-2c}$ . It can be seen that our expression is far more accurate, verifying that our expression is correct. Numerically we have estimated the next higher order term as

$$\delta \approx 4\sqrt{\pi c} e^{-2c} \left( 1 - \frac{7}{16c} - \frac{91}{2^9 c^2} - \frac{2657}{2^{13} c^3} \right). \quad (64)$$

The error for this expression is also shown in Fig. 9, and is even smaller.

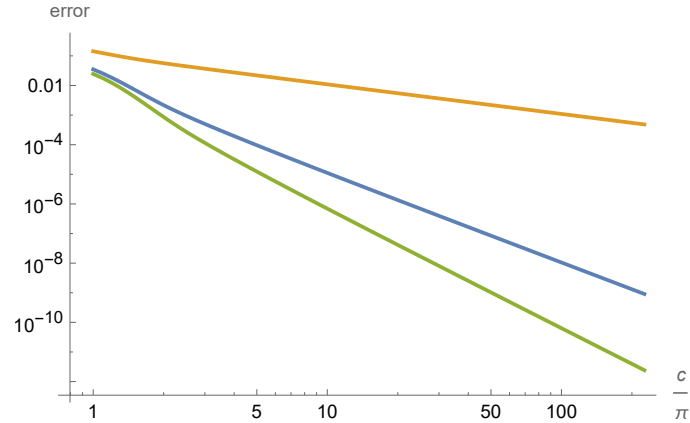


FIG. 9. The differences between the various asymptotic series and exact error, divided by  $4\sqrt{\pi c} e^{-2c}$ . Results for the series of Slepian [see Eq. (58)] are shown in orange, our series in Eq. (59) is given in blue, and our numerically estimated series in Eq. (64) is shown in green.

## V. THE SAMPLING METHOD

### A. Asymptotic approximations

When we are performing phase estimation on a block encoded Hamiltonian, the Hamiltonian is encoded as  $H/\lambda$ , and the eigenvalues of the corresponding qubitised operator are  $\pm e^{\pm i \arcsin(E_k/\lambda)}$  for eigenvalues  $E_k$  of  $H$ . This means

that restricting the error to  $\leq \epsilon/\lambda$  for the phase estimation means that the error in  $E_k$  is no more than  $\epsilon$ . In the following we denote the ground state energy by  $E_0$ , so we aim to have an estimate in the region  $[E_0 - \epsilon, E_0 + \epsilon]$ . Then the number of queries for phase estimation with either the Kaiser or Slepian window becomes, to leading order

$$\frac{\lambda}{2\epsilon} \ln(1/\delta) + \mathcal{O}((\lambda/\epsilon) \ln \ln(1/\delta)). \quad (65)$$

If you take the amplitude for the ground state to be  $\gamma$  and  $p = \gamma^2$ , then the probability of failing to have the ground state once in sampling the energies  $n$  times is  $(1-p)^n \approx e^{-pn}$ . If we want the probability of failing to be less than  $q$ , we would need to take

$$n \approx (1/p) \ln(1/q). \quad (66)$$

If we are taking the minimum result for the eigenvalue, then the probability of the error in that measurement result being outside the  $\epsilon$  interval is now bounded by  $n\delta$ . This suggests that we should divide  $\delta$  by  $n$  in order to obtain an appropriately bounded error probability. However, the importance of the errors is asymmetric. If the measurement corresponds to an excited eigenstate, but the measurement error yields an estimate of the energy that is exceptionally low, then that estimate could be taken to be the smallest out of all samples, yielding an inaccurate result. On the other hand, if the measurement error gives an estimate of the energy that is exceptionally high, then it is far less likely to be taken as the smallest estimate out of all samples. This means that the measurement errors in each direction need to be quantified differently.

To do this, let the confidence level for each individual estimate of the phase be  $1 - \delta$ , so the probability of error on *one side* is  $\delta/2$ . That is because we are choosing a measurement technique where the error distribution is symmetric. Given squared overlap with the ground state  $p$ , then for each estimate there is probability  $1 - p$  of it corresponding to the wrong eigenstate, and  $p\delta/2$  of it corresponding to the correct eigenstate but the true eigenvalue being below the confidence interval on the lower side (so the estimate is too large). We will group these possibilities together as a ‘high’ error, which has total probability  $1 - p(1 - \delta/2)$ . The probability of there being high errors on all  $n$  samples is then  $[1 - p(1 - \delta/2)]^n$ .

The probability of a ‘low’ error outside the confidence interval is  $\delta/2$  for any individual measurement, so the probability of any low error occurring in the  $n$  samples is  $1 - (1 - \delta/2)^n$ . Note that it is possible for measurements corresponding to excited states to give a low error that is still not below the desired confidence interval for the measurement of the ground state. The estimate above is not taking that into account, so is a fairly loose upper bound on the error. Our total upper bound on the error is then

$$[1 - p(1 - \delta/2)]^n + 1 - (1 - \delta/2)^n. \quad (67)$$

Given an allowable error  $q$  (confidence level  $1 - q$  in the final estimate), we can then solve for  $\delta$ . This approach results in the value of  $\delta$  being about twice what it would be if we did not take account of the asymmetry.

The overall complexity will then be approximately

$$\frac{n\lambda}{2\epsilon} \ln(1/\delta), \quad (68)$$

with  $\delta$  as chosen by solving for  $q$  equal to Eq. (67), and  $n$  chosen as at least  $(1/p) \ln(1/q)$ , resulting in a total complexity

$$\frac{\lambda}{2p\epsilon} \ln(1/q) \ln(1/\delta). \quad (69)$$

For any specific example we can tweak the value of  $n$  in order to minimise the overall complexity. As an example, let us consider  $\gamma = 0.1$  so  $p = 0.01$ , and require a 95% confidence interval so that  $q = 0.05$ . In that case we have the factor  $n \ln(1/\delta)/2$  as a function of  $n$  shown in Fig. 10. That is, this is the factor in the complexity that is multiplied by  $\lambda/\epsilon$  to give the overall complexity. Here it can be seen that the optimal value of  $n$  is 325, which is moderately above  $(1/p) \ln(1/q) \approx 300$ . Note that this choice of the optimal value of  $n$  is independent of  $\lambda$  and  $\epsilon$ .

We can further develop asymptotic approximations for the solutions in the case of small  $p, q$  to estimate the optimal values of  $n$  and  $\delta$ . First, we linearise Eq. (67) and set it equal to  $q$  to give

$$(1-p)^n + \frac{\delta n}{2} [1 + (1-p)^{n-1} p] \approx q. \quad (70)$$

Solving for  $\delta$  then gives

$$\delta \approx \frac{2}{n} \frac{q - (1-p)^n}{1 + (1-p)^{n-1} p}. \quad (71)$$

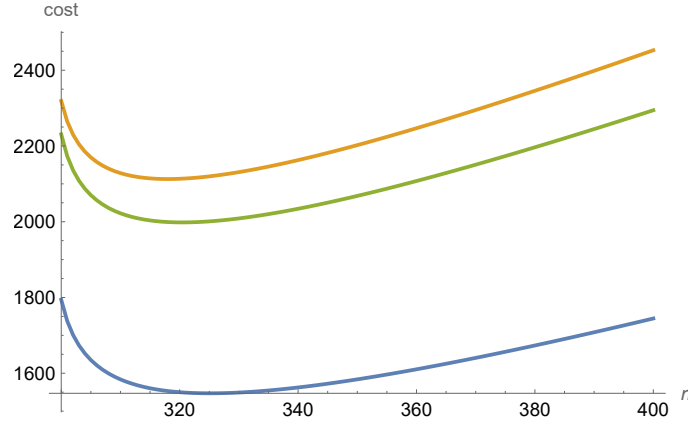


FIG. 10. The factor in the complexity as a function of  $n$  for  $p = 0.01$  and  $q = 0.05$ , with  $\delta$  obtained by solving for  $q$  equal to Eq. (67). The blue curve is the approximation  $n \ln(1/\delta)/2$ , the orange curve is the factor  $n\pi\sqrt{1+\alpha^2}$  for the Kaiser window, and the green curve is  $n\pi\sqrt{\Delta^2+\alpha^2}$  with  $\Delta = 0.3239$ .

If we expand  $\delta$  to second order, then we obtain the next term in the expansion

$$\delta \approx \frac{2}{n} \frac{q - (1-p)^n}{1 + (1-p)^{n-1}p} + \frac{(n-1)[1 - (1-p)^{n-2}]}{n^2[1 + (1-p)^{n-1}p]^3} (q - (1-p)^n)^2. \quad (72)$$

Up to terms that are smaller by factors of  $q$  or  $p$ , we can give

$$\frac{1}{\delta} \approx \frac{n}{2[q - (1-p)^n]} = \frac{n}{2(q - e^{-n\rho})}, \quad (73)$$

where  $\rho = -\ln(1-p) \approx p$ .

Next, given the solution for  $\delta$  the task is to choose  $n$  to minimise  $n \ln(1/\delta)$ . Let us first consider this expression with only the leading order in the solution for  $\delta$ , which is (ignoring the factor of 2 for the moment)

$$n \ln \left( \frac{n}{2(q - e^{-n\rho})} \right). \quad (74)$$

The higher-order terms in the solution for  $\delta$  will result in corrections that are at least a factor of  $q$  smaller, so can be safely ignored in the following analysis where the terms in the expansions are significantly larger. Taking the derivative with respect to  $n$  gives

$$1 - \frac{e^{-n\rho}n\rho}{q - e^{-n\rho}} + \ln \left( \frac{n}{2(q - e^{-n\rho})} \right). \quad (75)$$

Setting this to zero and rearranging gives

$$h = g - \ln g, \quad (76)$$

where

$$g = \frac{e^{-n\rho}n\rho}{q - e^{-n\rho}}, \quad (77)$$

$$h = 1 + n\rho - \ln(2\rho). \quad (78)$$

We can then solve for  $g$  as a series in  $h$  as

$$g \approx h + \ln h + \frac{\log h}{h} + \frac{2 \ln h - \ln^2 h}{2h^2} + \frac{6 \ln h - 9 \ln^2 h + 2 \ln^3 h}{6h^3} + \mathcal{O} \left( \frac{\ln^4 h}{h^4} \right). \quad (79)$$

Now  $q$  can be given in terms of  $g$  as

$$q = e^{-n\rho} + \frac{e^{-n\rho}n\rho}{g}. \quad (80)$$

We can then rewrite this as

$$\ln(2/q) = n\rho - \ln\left(\frac{1}{2} + \frac{n\rho}{2g}\right). \quad (81)$$

By substituting a series for  $g$  in terms of  $h$ , and inverting to obtain a series for  $n\rho$ , we obtain

$$n\rho = \ln(2/q) - \frac{Q+1}{2\ln(2/q)} \left(1 - \frac{3Q-1}{4\ln(2/q)} - \frac{5+4Q-7Q^2}{12\ln^2(2/q)}\right) + \mathcal{O}\left(\frac{Q^4}{\ln^4(2/q)}\right), \quad (82)$$

$$Q := \ln(\ln(2/q)/2\rho). \quad (83)$$

The difficulty with using this expression is that the successive terms are not smaller if  $Q \sim \ln(2/q)$ . That will be the case if  $q > p$ . This expression does give accurate results if  $q$  is small compared to  $p$ .

A more accurate approximation may be given by

$$n\rho = \ln(2/q) - \frac{Q+1}{2\ln(2/q)} \left(1 + \frac{3Q-1}{4\ln(2/q)} + \frac{23-2Q-Q^2}{48\ln^2(2/q)}\right)^{-1} + \mathcal{O}\left(\frac{Q^4}{\ln^4(2/q)}\right). \quad (84)$$

Using this expression with  $p = 0.01$  and  $p = 0.05$  gives  $n \approx 325$ , close to the true value. This series then gives the leading-order terms for  $n\ln(1/\delta)/2$  as

$$\frac{n}{2} \ln(1/\delta) = \frac{1}{2\rho} \ln(2/q) \ln\left[\frac{\ln(2/q)}{\rho q}\right] - \frac{(1+Q)^2}{8\rho \ln(2/q)} + \frac{(Q-1)(1+Q)^2}{16\rho \ln^2(2/q)} + \frac{(1+Q)^2(5+10Q-7Q^2)}{192\rho \ln^3(2/q)} + \mathcal{O}\left(\frac{Q^5}{\rho \ln^4(2/q)}\right). \quad (85)$$

For this example the estimated value using the terms shown is 1525, within 1.5% of the true value of 1547. Using only the leading term gives 1634, still within 6% of the correct value. Note also that using only the leading order term for  $\delta$  in estimating the cost affects the result by very little, less than 0.06% for this example. This justifies omitting higher-order terms for  $\delta$  in the above analysis. In contrast, using the approximation  $\rho \approx p$  affects the result by about 0.6%, so it is useful to make that correction.

## B. Real cost for window functions

Next we consider the actual cost for the Kaiser and prolate spheroidal window functions, rather than just the asymptotic approximation. We can explicitly integrate the probability distribution for the error in the Kaiser window. That gives a factor of  $n\pi\sqrt{1+\alpha^2}$  rather than  $n\ln(1/\delta)/2$ , with  $\alpha$  solved to give error  $\delta$ . The resulting factor is also shown in Fig. 10. The factor has been increased by about 37% above that for the asymptotic approximation given above, from about 1547 to 2113. It is also possible to adjust the cutoff used in the Kaiser window to  $\sqrt{\Delta^2 + \alpha^2}$  for  $\Delta \neq 1$ , which can give improved results. The result for  $\Delta = 0.3239$  gives approximately the optimal result, and is shown in Fig. 10. Now the constant factor is only increased by about 29% to 1998. The accuracy of the asymptotic approximation is better for smaller  $q$ , so  $q = 0.01$  results in the cost being about 25% above the asymptotic value, but it would take very small  $q$  for the approximation to be accurate.

The prolate spheroidal window further reduces the cost, but only by a very small amount. The curve is indistinguishable from that for optimised  $\Delta$  in Fig. 10, so is not shown separately. In this case we find the constant factor is reduced very slightly to 1997 with  $n = 320$ , a reduction of only 0.06%. This shows that it is possible to accurately approximate results for the optimal window using the Kaiser window and adjusting  $\Delta$ .

## C. Contribution to cost from excited states

We can give a tighter bound on the cost by more accurately accounting for the contribution to the error from excited states. In practice, there will be a small contribution to the probability of low estimates of the ground state energy  $E_0$  from excited states. When they are distant from  $E_0$ , there is very little probability of them yielding an estimate of the eigenvalue below the desired confidence interval, and if they are close to  $E_0$  they will also increase the probability of having a result within the desired confidence interval.

To gauge the effect, let us consider just a single excited state with energy  $\beta\epsilon$  above the ground state. That is, when  $\beta < 1$  the energy is actually within the desired confidence interval for estimating the ground state. It can be shown that the error is maximised for a single excited state, so our analysis for a single excited state is sufficient to bound

the worst case for any spectrum of excited states; see Appendix D. We denote by  $\delta_1$  the probability for the estimate above the confidence interval for the excited state, and  $\delta_2$  for the estimate below the desired confidence interval. It will be expected that  $\delta < \delta_1 + \delta_2$ , and in typical cases where the eigenvalue is *much* higher than the ground value, there will be  $\delta_1 \sim 1$  and  $\delta_2 \sim 0$ .

We can then replace  $1 - p$  for the probability of error due to the incorrect eigenstate being obtained with  $(1 - p)\delta_1$ ; that is, the incorrect eigenstate is obtained *and* the estimate is too high. Then the probability of all  $n$  samples being too high is

$$[(1 - p)\delta_1 + p\delta/2]^n. \quad (86)$$

Then the probability of a single sample being too low is  $p\delta/2 + (1 - p)\delta_2$ , so the probability of *any* of the samples being too low is  $1 - \{1 - [p\delta/2 + (1 - p)\delta_2]\}^n$ . Adding together these two probabilities of error then gives

$$P_{\text{err}} = [p\delta/2 + (1 - p)\delta_1]^n + 1 - \{1 - [p\delta/2 + (1 - p)\delta_2]\}^n. \quad (87)$$

This expression can be expected to be smaller than that given before, because  $p\delta/2 + (1 - p)\delta_1$  will be smaller than  $p\delta/2 + (1 - p)$ .

In numerical testing, we choose values of  $\alpha, \Delta, n$  *without* knowing  $\beta$  (the excited state energy), so the above error needs to be no greater than  $q$  for all choices of  $\beta$ . We find that for  $p = 0.01$ ,  $q = 0.05$ , we can choose  $\alpha = 1.70116$ ,  $\Delta = 0.074476$ , and  $n = 309$ . These values minimise the constant factor while keeping the error below  $q$ ; see Fig. 11. The constant factor is approximately 1673, so it is significantly smaller than the results not taking into account this factor, and similar to the result with the very simple asymptotic approximation we gave first. In testing with multiple excited state eigenvalues, the results are no worse than with just one excited state eigenvalue, as predicted.

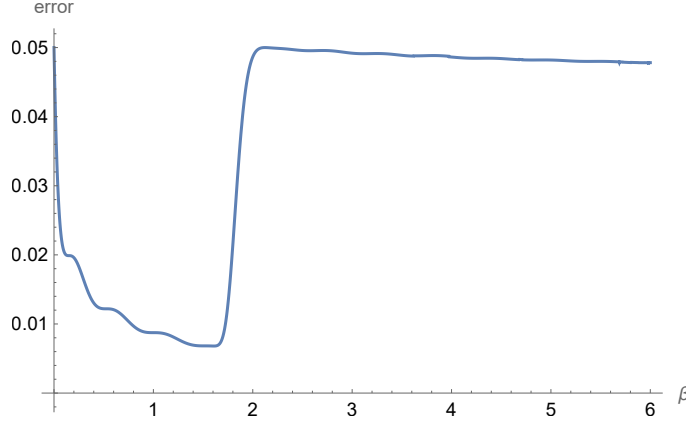


FIG. 11. The error according to Eq. (87) as a function of  $\beta$  with  $p = 0.01$ ,  $\alpha = 1.70116$ ,  $\Delta = 0.074476$ , and  $n = 309$ . The error does not exceed  $q = 0.05$ .

An interesting feature of the results is that the error is larger for larger  $\beta$  (around 2 and above), then is smaller for small  $\beta$ , but with a spike near  $\beta = 0$ . The reason for this is that, as  $\beta$  is reduced below 2, there is greater overlap between the distribution of measurement results obtained for the excited state and the desired interval  $[E_0 - \epsilon, E_0 + \epsilon]$  for estimates of  $E_0$ . That is, both of these are of half-width  $\epsilon$ , so reducing the gap below  $\beta\epsilon$  means the two regions overlap. This means that measurements of energy for the excited state being within the desired region is reducing the overall error probability. It reduces quite quickly, because we are taking the minimum of the measurement results, and with many samples there is a high probability of a measurement result on the excited state being around  $\epsilon$  below its true energy.

But then there is a separate spike for the error probability as the gap closes to zero. This is because the measurements of energy on the excited state now have a significant probability of being *below* the desired confidence interval. To bound the error by  $q$  we need to ensure the error for both  $\beta = 0$  and  $\beta$  around 2 is bound by  $q$ . We find that the best results (in terms of the lowest constant factor) tends to be those where the error reaches  $q$  for both. This is illustrated in Fig. 12. It shows that the probability of errors more than  $\epsilon$  above the true ground state rapidly approaches zero as  $\beta$  is reduced below 2, and the probability of errors that are too low gradually increases. This behaviour is for small  $p$ , whereas for  $p$  close to 1 the error is less than  $q$  for  $\beta = 0$ .

It turns out that using the prolate spheroidal wavefunction now gives *worse* results. For this example we find that the performance is optimised for  $n = 318$ ,  $\alpha = 1.71229$ , where the constant factor is about 1711, or 2% *worse* than



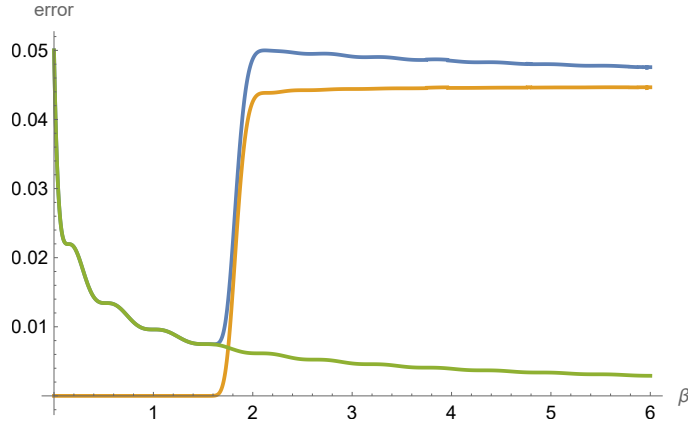


FIG. 12. The error according to Eq. (87) as a function of  $\beta$  with  $p = 0.01$ ,  $\alpha = 1.70116$ ,  $\Delta = 0.074476$ , and  $n = 309$ . The green line is the probability of an error below the desired interval, the orange line is for errors above the interval, and blue is the sum of the two.

for the Kaiser window. To see why, consider the same parameters as for the Kaiser window, with  $\beta = 2.12103$ . We then find that most of the probabilities are the same, but  $\delta_2$  is significantly lower for the Kaiser window. It is about 0.0000184942, versus  $\delta_2 \approx 0.0000336569$  for the prolate spheroidal window, which is about 82% larger.

This is very significant in this case with small  $p$ , because it contributes to the chance of the excited state yielding an estimate of the energy that is too *low*. In this case, with the prolate spheroidal wavefunction, the probability of an estimate that is too low with the excited state is only about 1/5 that for the ground state. Since the excited state is about 100 times more common in this example, that is a significant contributor to the error. In contrast, the Kaiser window more strongly suppresses the tails, so estimates that are too low coming from the excited state are less of a problem. Again the behaviour is different for  $p$  close to 1, where the prolate spheroidal window provides better performance. That is because there are few repetitions, so the effect of the contribution to the error from  $\delta_2$  being amplified by repetitions is less.

## VI. THE BINARY SEARCH APPROACH

In the previous sections we have focused on estimating the ground state energy by directly reading off the energy samples from the quantum phase estimation algorithm and taking the minimum among these samples. As shown in Eq. (66), the number of samples needed scales as  $\tilde{\mathcal{O}}(1/p)$ . This is however not the optimal scaling. As shown in [26, Theorem 8], one can improve the dependence on  $p$  to  $\tilde{\mathcal{O}}(1/\sqrt{p})$ . In this section we will propose a method based on [26] and incorporating the window functions to reduce the resources needed in practical implementation. We also show numerically that this method is beneficial when  $p < 10^{-3}$ . We will refer to this method as *the binary search approach* henceforth.

### A. From the fuzzy bisection problem to amplitude estimation

From the previous sections, we can see that in a single run of the QPE algorithm, in order to ensure that the phase error is below  $\eta$  with probability at least  $1 - \delta$ , the query complexity is given by a function  $Q(\eta, \delta)$ , which in the case of the Kaiser or Slepian window, scales as

$$Q(\eta, \delta) \approx \frac{1}{2\eta} \ln(1/\delta) \quad (88)$$

for small  $\eta$  and  $\delta$ , according to Eq. (52). In this notation we use  $\eta$  for the phase estimation error to distinguish it from the error  $\epsilon$  for estimation of eigenvalues. We will use this expression in our computation of the asymptotic complexity of the binary search approach, but in numerics we will numerically compute the function  $Q(\eta, \delta)$  in a more accurate manner.

In the binary search process in Ref. [26], we gradually shrink an interval  $[\lambda_L, \lambda_R]$  in which the ground state energy is located. In particular, in the last search step, in order to estimate  $\lambda_0$  to  $\epsilon$  precision, we have  $\lambda_L, \lambda_R$  such that

$\lambda_L \leq \lambda_0 \leq \lambda_R$ , with  $\lambda_R - \lambda_L = 3\epsilon$ . We want to distinguish between two cases:

$$\lambda_0 > \frac{1}{3}\lambda_L + \frac{2}{3}\lambda_R, \quad \text{or} \quad \lambda_0 < \frac{2}{3}\lambda_L + \frac{1}{3}\lambda_R. \quad (89)$$

If  $\frac{2}{3}\lambda_L + \frac{1}{3}\lambda_R \leq \lambda_0 \leq \frac{1}{3}\lambda_L + \frac{2}{3}\lambda_R$  then we can output anything. Solving the problem of distinguishing the cases will give us an interval  $[\lambda_L^{\text{new}}, \lambda_R^{\text{new}}] \ni \lambda_0$  of size  $2\epsilon$ . That interval corresponds to an estimate of  $\lambda_0$  up to  $\epsilon$  error. We call this problem the *fuzzy bisection problem*.

Because  $\arccos$  is monotonically decreasing in  $[-1, 1]$ , we only need to distinguish between

$$\arccos\left(\frac{\lambda_0}{\lambda}\right) < \arccos\left(\frac{\lambda_L + 2\lambda_R}{3\lambda}\right), \quad \text{or} \quad \arccos\left(\frac{\lambda_0}{\lambda}\right) > \arccos\left(\frac{2\lambda_L + \lambda_R}{3\lambda}\right). \quad (90)$$

We then perform QPE with a Kaiser or Slepian window on  $W$ . We denote the phase output by  $\hat{\phi} \in [-\pi, \pi]$ . We want the phase error to be at most

$$\frac{\epsilon}{2\lambda} = \frac{\lambda_R - \lambda_L}{6\lambda} \leq \frac{1}{2} \left( \arccos\left(\frac{2\lambda_L + \lambda_R}{3\lambda}\right) - \arccos\left(\frac{\lambda_L + 2\lambda_R}{3\lambda}\right) \right) \quad (91)$$

with probability at least  $1 - \delta_1$ .

If  $\arccos\left(\frac{\lambda_0}{\lambda}\right) < \arccos\left(\frac{\lambda_L + 2\lambda_R}{3\lambda}\right)$ , then all eigenvalues  $e^{\pm i \arccos(\lambda_k/\lambda)}$  of  $W$  satisfy  $\arccos\left(\frac{\lambda_k}{\lambda}\right) < \arccos\left(\frac{\lambda_L + 2\lambda_R}{3\lambda}\right)$ . Therefore if

$$|\hat{\phi}| > \bar{\phi} := \frac{1}{2} \left( \arccos\left(\frac{2\lambda_L + \lambda_R}{3\lambda}\right) + \arccos\left(\frac{\lambda_L + 2\lambda_R}{3\lambda}\right) \right), \quad (92)$$

then a phase error larger than

$$\bar{\phi} - \arccos\left(\frac{\lambda_L + 2\lambda_R}{3\lambda}\right) \geq \frac{\epsilon}{2\lambda} \quad (93)$$

must have occurred. This event has probability at most  $\delta_1$ . Consequently

$$\Pr[|\hat{\phi}| > \bar{\phi}] \leq \delta_1. \quad (94)$$

If  $\arccos\left(\frac{\lambda_0}{\lambda}\right) > \arccos\left(\frac{2\lambda_L + \lambda_R}{3\lambda}\right)$ , suppose  $\hat{\phi}$  comes from eigenvalues  $e^{\pm i \arccos(\lambda_0/\lambda)}$ , then it will satisfy  $|\hat{\phi}| > \bar{\phi}$  with probability at least  $1 - \epsilon_1$ . This is because

$$\arccos\left(\frac{2\lambda_L + \lambda_R}{3\lambda}\right) - \bar{\phi} \geq \frac{\epsilon}{2\lambda}. \quad (95)$$

The phase  $\hat{\phi}$  comes from eigenvalues  $e^{\pm i \arccos(\lambda_0/\lambda)}$  with probability at least  $p$ . Therefore

$$\Pr[|\hat{\phi}| > \bar{\phi}] \geq p(1 - \delta_1). \quad (96)$$

We therefore only need to distinguish between two cases in Eq. (94) and Eq. (96). This is an amplitude estimation problem. We define

$$\gamma_1 = \sqrt{\delta_1}, \quad \gamma_2 = \sqrt{p(1 - \delta_1)}, \quad (97)$$

and aim to distinguish the cases where the amplitude corresponding to  $|\hat{\phi}| > \bar{\phi}$  is at most  $\gamma_1$  or at least  $\gamma_2$ . We choose the parameters so that  $\gamma_1 < \gamma_2$ . To generate a single sample of  $\hat{\phi}$  we need to run a coherent QPE circuit that involves

$$d_1 = Q(\eta, \delta_1) \quad (98)$$

queries to  $W$ , where  $\eta = \epsilon/(2\lambda)$  is the allowed phase error.

### B. Amplitude estimation with the Kaiser window

From the previous section we can see that to solve the fuzzy bisection problem, it suffices to estimate

$$A = \sqrt{\Pr[|\hat{\phi}| > \bar{\phi}]}.$$
 (99)

We will do so using amplitude amplification. If this amplitude can be estimated with error at most  $(\gamma_2 - \gamma_1)/2$ , then we will be able to distinguish between  $A > \gamma_2$  and  $A < \gamma_1$ . In amplitude estimation, we construct a walk operator  $\mathcal{W}$  using the QPE circuit (two applications of it), such that

$$\mathcal{W}|\Phi^\pm\rangle = e^{\pm i2 \arcsin(A)} |\Phi^\pm\rangle,$$
 (100)

and  $|\Phi\rangle = \frac{1}{2}(|\Phi^+\rangle + |\Phi^-\rangle)$  can be prepared using the QPE circuit.

With this walk operator  $\mathcal{W}$ , we can then run QPE with the Kaiser or Slepian window to estimate  $\arcsin(A)$ . We only need to estimate  $A$  to precision  $(\gamma_2 - \gamma_1)/2$ , which means it suffices to estimate the phase  $2 \arcsin(A)$  to precision  $\gamma_2 - \gamma_1$ . To do this with probability at least  $1 - \delta_2$  requires running  $\mathcal{W}$   $d_2$  times, where

$$d_2 = Q(\gamma_2 - \gamma_1, \delta_2).$$
 (101)

For the last search step, we need to use  $W$  for a total of

$$d_1(2d_2 + 1) = Q\left(\frac{\epsilon}{2\lambda}, \delta_1\right)(2Q(\gamma_2 - \gamma_1, \delta_2) + 1)$$
 (102)

times, which in the context of the Kaiser window and using Eq. (65) is

$$d_1(2d_2 + 1) \approx 2d_1d_2 \approx \frac{\lambda}{(\gamma_2 - \gamma_1)\epsilon} \ln(1/\delta_1) \ln(1/\delta_2)$$
 (103)

times, up to the leading order. Note that in  $2d_2 + 1$  the  $+1$  comes from preparing the initial state  $\mathcal{W}|0\rangle$  for amplitude estimation (see [42, Figure 1]). The number of times we need to use  $U_{\text{init}}$  is  $2d_2 + 1$ , which in the context of the Kaiser window is

$$2d_2 + 1 \approx 2d_2 = \frac{1}{(\gamma_2 - \gamma_1)} \ln(1/\delta_2).$$
 (104)

The value of  $\delta_1$  can be chosen to minimise the cost as follows. Using the expressions  $\gamma_1 = \sqrt{\delta_1}$ ,  $\gamma_2 = \sqrt{p(1 - \delta_1)}$  we have a factor in the complexity

$$\frac{\ln(1/\delta_1)}{\sqrt{p(1 - \delta_1)} - \sqrt{\delta_1}}.$$
 (105)

Approximating  $\sqrt{1 - \delta_1} \approx 1$  and using  $x = \sqrt{\delta_1}$  this factor is approximately proportional to

$$\frac{\ln(1/x)}{\sqrt{p} - x}.$$
 (106)

Taking the derivative with respect to  $x$  then yields

$$\frac{-1/x}{\sqrt{p} - x} + \frac{\ln(1/x)}{(\sqrt{p} - x)^2}.$$
 (107)

For this to be zero, we should have

$$x = \frac{\sqrt{p} - x}{\ln(1/x)}.$$
 (108)

Starting with  $x = \sqrt{p}/\ln(1/\gamma)$  then iterating  $x \mapsto (\sqrt{p} - x)/\ln(1/x)$  quickly gives the solution.

### C. Query complexity of all search steps

Previously we focused on the cost of the last search step. Here we account for the costs in all search steps: the binary search terminates in  $L = \lceil \log_{3/2}(\lambda/\epsilon) \rceil$  steps. If we want a final success probability of  $q$ , we need  $\delta_2 = q/L$ . For the number of queries to  $W$ , we observe that each search step requires  $2/3$  of the resolution of the next step, and therefore the query complexity is also  $2/3$  of that of the last step. Therefore the total number of queries to  $W$  is

$$\begin{aligned} & Q\left(\frac{\epsilon}{2\lambda}, \delta_1\right) (2Q(\gamma_2 - \gamma_1, \delta_2) + 1) \left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \dots\right) \\ & \leq 3Q\left(\frac{\epsilon}{2\lambda}, \delta_1\right) (2Q(\gamma_2 - \gamma_1, \delta_2) + 1). \end{aligned} \quad (109)$$

For the Kaiser window the total number of queries needed is then

$$\frac{3\lambda}{(\gamma_2 - \gamma_1)\epsilon} \ln(1/\delta_1) \ln(1/\delta_2). \quad (110)$$

The total number of queries to  $U_{\text{init}}$  is the same for each search step. Therefore it is

$$\frac{L}{\gamma_2 - \gamma_1} \ln(1/\delta_2). \quad (111)$$

We recall that

$$L \approx \log_{3/2}(\lambda/\epsilon), \quad \delta_2 = q/L. \quad (112)$$

At the end of the previous section we have discussed how to choose  $\delta_1$  by solving an optimization problem, but here to get a concise expression we will choose a sub-optimal  $\delta_1$ , which does not have much effect on the final cost. Here  $\delta_1$  is chosen to be

$$\delta_1 = p/16, \quad (113)$$

and then

$$\gamma_2 = \sqrt{p(1 - \delta_1)} = \sqrt{p(1 - p/16)}, \quad \gamma_1 = \sqrt{\delta_1} = \sqrt{p}/4. \quad (114)$$

Consequently

$$\gamma_2 - \gamma_1 \approx (3/4)\sqrt{p}. \quad (115)$$

Substituting these values into Eq. (110) and Eq. (111), we can see that in order to estimate the ground state energy to precision  $\epsilon$  with probability at least  $1 - q$ , with an initial squared overlap of  $p$ , we need to use  $W$

$$\frac{8\lambda}{\sqrt{p}\epsilon} \ln\left(\frac{4}{\sqrt{p}}\right) \ln\left(\frac{\log_{3/2}(\lambda/\epsilon)}{q}\right) \quad (116)$$

times, and we need to use  $U_{\text{init}}$

$$\frac{4 \log_{3/2}(\lambda/\epsilon)}{3\sqrt{p}} \ln\left(\frac{\log_{3/2}(\lambda/\epsilon)}{q}\right) \quad (117)$$

times.

### D. Optimizing the shrinking factor

In the binary search algorithm we discussed above, we shrink the interval in which  $\lambda_0$  is located by  $2/3$  in each iteration. This is however not the optimal shrinking factor. If we shrink by a factor of  $\omega$  instead of  $2/3$  at each search

step, then using the leading order in the asymptotic expression for the Kaiser and Slepian windows in Eq. (52), we need to use  $W$

$$\underbrace{\frac{1}{1-\omega}}_{\text{search steps}} \times \underbrace{\frac{\omega}{2(2\omega-1)} \frac{\lambda}{\epsilon} \ln\left(\frac{16}{p}\right)}_{\text{QPE}} \times \underbrace{\frac{4}{3\sqrt{p}} \ln\left(\frac{\log_{1/\omega}(\lambda/\epsilon)}{q}\right)}_{\text{amplitude estimation}} = \frac{4\omega}{3(2\omega-1)(1-\omega)} \frac{\lambda}{\sqrt{p}\epsilon} \ln\left(\frac{4}{\sqrt{p}}\right) \ln\left(\frac{\log_{1/\omega}(\lambda/\epsilon)}{q}\right) \quad (118)$$

times.

We will explain how we arrived at the above expressions. For each binary search step, the two cases we wish to distinguish are now

$$\lambda_0 > (1-\omega)\lambda_L + \omega\lambda_R, \quad \text{or} \quad \lambda_0 < \omega\lambda_L + (1-\omega)\lambda_R. \quad (119)$$

If we have eliminated the first case, then the value of  $\lambda_R$  is mapped as

$$\lambda_R \mapsto (1-\omega)\lambda_L + \omega\lambda_R, \quad (120)$$

with  $\lambda_L$  unchanged, so the range of values is mapped as

$$\lambda_R - \lambda_L \mapsto (1-\omega)\lambda_L + \omega\lambda_R - \lambda_L = \omega(\lambda_R - \lambda_L). \quad (121)$$

The reduction in the range is equivalent if we eliminate  $\lambda_0 < \omega\lambda_L + (1-\omega)\lambda_R$ .

We now need phase error at most

$$\frac{(2\omega-1)(\lambda_R - \lambda_L)}{2\lambda} \leq \frac{1}{2} \left[ \arccos\left(\frac{\omega\lambda_L + (1-\omega)\lambda_R}{\lambda}\right) - \arccos\left(\frac{(1-\omega)\lambda_L + \omega\lambda_R}{\lambda}\right) \right]. \quad (122)$$

At each step the range is shrunk by  $\omega$ , and we start with  $\lambda_R - \lambda_L = 2\lambda$ , so at step  $j$  the error is

$$(2\omega-1)\omega^{j-1}. \quad (123)$$

The number of queries for the QPE circuit is then approximately, using Eq. (65),

$$\frac{1}{2(2\omega-1)} (1/\omega)^{j-1} \ln(1/\delta_1). \quad (124)$$

Now the number of steps  $L$  is calculated so that

$$(1/\omega)^L \lambda = \epsilon. \quad (125)$$

If we sum the cost above over  $j = 1$  to  $L$ , then we obtain

$$\frac{1}{2(2\omega-1)} \times \frac{(1/\omega)^L - 1}{1/\omega - 1} \ln(1/\delta_1) \approx \frac{\omega}{2(2\omega-1)(1-\omega)} \frac{\lambda}{\epsilon} \ln(1/\delta_1). \quad (126)$$

To minimise the cost we therefore aim to maximise  $(2\omega-1)(1/\omega-1)$ . That has a derivative of  $1/\omega^2 - 2$ , so the turning point is  $\omega = 1/\sqrt{2}$ .

Note that an extra factor needs to be taken into account for amplitude estimation. This is independent of the choice of  $\omega$  and therefore identical to the case we studied before for  $\omega = 2/3$ , which is given in Eq. (110). With the parameters  $\gamma_1$  and  $\gamma_2$  given in Eq. (114), we thus arrive at the expressions in Eq. (118) by multiplying this extra factor by Eq. (126).

Optimizing  $\omega$  for (118) outside the  $\ln \ln$  we have  $\omega = 1/\sqrt{2}$ , and the number of queries are roughly

$$\frac{7.77\lambda}{\sqrt{p}\epsilon} \ln\left(\frac{4}{\sqrt{p}}\right) \ln\left(\frac{\log_{\sqrt{2}}(\lambda/\epsilon)}{q}\right). \quad (127)$$

We need to use  $U_{\text{init}}$

$$\underbrace{\log_{1/\omega}(\lambda/\epsilon)}_{\text{search steps}} \times \underbrace{\frac{4}{3\sqrt{p}} \ln\left(\frac{\log_{1/\omega}(\lambda/\epsilon)}{q}\right)}_{\text{amplitude estimation}}$$

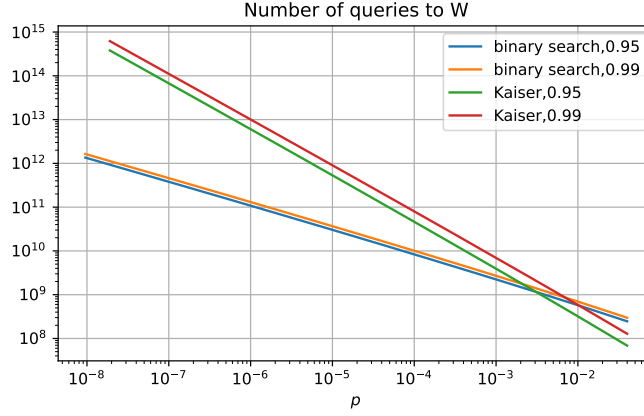


FIG. 13. Number of queries to  $W$  as a function of the squared overlap  $p$  for FeMoco with  $\lambda = 306$  and  $\epsilon = 0.0016$ , using both the binary search method discussed in this section and pure QPE+Kaiser window. At  $p = 0.01$ , the number of queries using the binary search method are 571 million and 691 million for 95% confidence and 99% confidence respectively, compared to 320 million and 587 million using pure QPE+Kaiser window. The binary sampling cost is affected very little by whether the Kaiser or optimal window is used for phase estimation. For the 95% and 99% confidence intervals the cost is only reduced by 0.04% and 0.06%, respectively. The value of  $\omega$  is chosen to be  $1/\sqrt{2}$ .

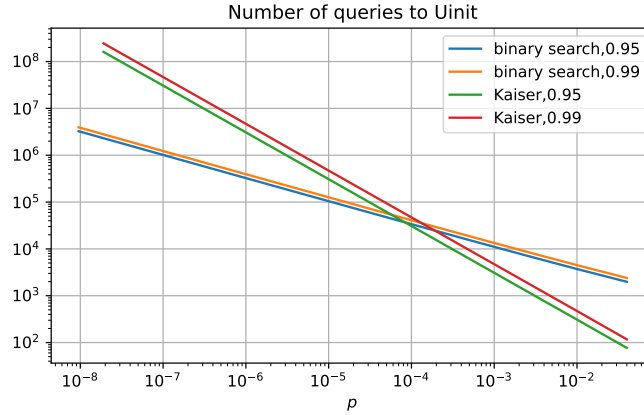


FIG. 14. Number of queries to  $U_{\text{init}}$  as a function of the squared overlap  $p$  for FeMoco with  $\lambda = 306$  and  $\epsilon = 0.0016$ , using both the binary search method discussed in this section and pure QPE+Kaiser window. At  $p = 0.01$ , the number of queries using the binary search method are 3683 and 4457 for 95% confidence and 99% confidence respectively, compared to 309 and 472 using pure QPE+Kaiser window. The value of  $\omega$  is chosen to be  $1/\sqrt{2}$ .

$$= \frac{4 \log_{1/\omega}(\lambda/\epsilon)}{3\sqrt{p}} \ln \left( \frac{\log_{1/\omega}(\lambda/\epsilon)}{q} \right) \quad (128)$$

times, which is monotonically increasing with respect to  $\omega$ .

In Figs 13 and 14 we numerically compute the query complexities of the binary search method and the direct sampling approach, both based on the Kaiser window. The results suggest that a crossover of the query complexity takes place between  $p = 10^{-3}$  and  $p = 10^{-2}$  for 95% and 99% confidence levels. This agrees well with the estimated crossover of around  $p \sim 0.003$  based on the constant factors in the scaling.

## VII. TOTAL RESOURCES FOR GROUND STATE ENERGY ESTIMATION IN CHEMICAL SYSTEMS

As previously discussed, a key quantity in the cost of applying QPE to ground state energy estimation is the overlap of the initial input state with the ground state wavefunction. Prior work investigated overlaps for simple

small molecules [3] and found high overlaps with product state wavefunctions substantially above the low  $p$  settings discussed in this work. More recently, it was shown that in realistic examples of more complicated molecules, product states can have a small overlap, while adiabatic state preparation from the lowest energy mean-field state can be more costly than the phase estimation itself, motivating the development of more sophisticated state preparation protocols [12]. However, determining an overlap, or a faithful estimate of it, in problems where classical algorithms have difficulty computing an accurate ground state (and thus which are of most interest for quantum applications) is by definition challenging.

In this section we describe an extrapolation protocol that allows us to estimate the overlap with the ground state by extrapolating with respect to the bond dimensions of two MPS wavefunctions. Although this extrapolation does not provide any rigorous guarantees, we provide benchmark data to support the procedure; see Appendix G. In addition, in the most challenging systems such as FeMoco, it is not possible to classically determine MPS wavefunctions with sufficiently large bond dimension to distinguish between the ground state and nearby excited states. In this case, under the assumption that the MPS we are preparing has good overlap with some low-energy eigenstate, we provide an estimation of the overlap with this eigenstate, and QPE can then be used to resolve the energies of the different eigenstates associated with the different MPS initial states. Leveraging these estimates, we account for the full resources needed for resolving the energy landscape of competing spin structures in FeMoco.

### A. Model generation

In this study, we employed active space models for Fe-S simulations suggested in previous studies for the  $\text{Fe}_2\text{S}_2$ ,  $\text{Fe}_4\text{S}_4$ , and FeMoco iron-sulfur systems [29, 43]. The active spaces for the  $\text{Fe}_2\text{S}_2$  and  $\text{Fe}_4\text{S}_4$  models consist of Fe-3d, S-3p, and  $\sigma$ -bonding orbitals between the Fe and thiolate ligands, defined as complete active spaces CAS(30e,20o) for  $2\text{Fe(III)}$ , CAS(54e,36o) for  $2\text{Fe(III)}2\text{Fe(II)}$ , and CAS(52e,36o) for  $4\text{Fe(III)}$ . In this notation CAS( $ne,mo$ ) indicates the complete active space with  $n$  being the number of electrons and  $m$  the spatial orbitals in the active space (with e and o labelling electrons and spatial orbitals). The active space model of FeMoco also includes additional Mo-4d and central C-2s and C-2p orbitals, resulting in a CAS(113e,76o) for  $4\text{Fe(III)}3\text{Fe(II)}\text{Mo(III)}$ .

The energy landscape of ansatz approximations to FeMoco ground-state wavefunctions with total spin  $S = 3/2$  is characterized by numerous local electronic minima. Therefore, obtaining the correct ground state requires starting from a good initial guess. To achieve this, we used a density matrix renormalization group (DMRG) initialization procedure as described in earlier studies [28, 29] within the implementation in BLOCK2 [44, 45]. For the initial guess in the spin-adapted DMRG calculations, we first performed a spin-projected MPS calculation, initiated by a spin-projected broken-symmetry determinant. We explored 35 different spin-projected determinants corresponding to the broken-symmetry configuration of  $\{2\text{Fe(III)}\uparrow, 2\text{Fe(III)}\downarrow, 2\text{Fe(II)}\uparrow, \text{Fe(II)}\downarrow, \text{Mo(III)}\downarrow\}$ , which were previously studied using broken-symmetry density functional theory in Ref. [46]. The resulting spin-projected MPS was then optimized up to a bond dimension of 50.

These initial MPSs were subsequently optimized using spin-adapted DMRG calculations, with the bond dimension increased to 2000. Out of these 35 MPSs, we selected the three lowest-energy states at this bond dimension as example initial states, denoted MPS1, MPS2, and MPS3. The character of the spin-couplings of these states, as represented by the spin-projected determinants used to initialize them, is shown in Fig. 15(a). The three MPSs were further optimized, increasing the bond dimension to 7000 for MPS1 and 4000 for MPS2 and MPS3. Figure 15(b) shows the spin correlation matrix  $\langle S_A \cdot S_B \rangle$  between metal centers  $\{A, B\}$  of these further optimized MPSs, which is defined by

$$S_A \cdot S_B = \sum_{\mu \in \{x,y,z\}} S_A^\mu S_B^\mu \quad \text{with} \quad (129)$$

$$S_A^\mu = \sum_{p \in A} s_p^\mu, \quad (130)$$

$$s_p^x = \frac{1}{2} \left( a_{p\uparrow}^\dagger a_{p\downarrow} + a_{p\downarrow}^\dagger a_{p\uparrow} \right), \quad (131)$$

$$s_p^y = \frac{1}{2i} \left( a_{p\uparrow}^\dagger a_{p\downarrow} - a_{p\downarrow}^\dagger a_{p\uparrow} \right), \quad (132)$$

$$s_p^z = \frac{1}{2} \left( a_{p\uparrow}^\dagger a_{p\uparrow} - a_{p\downarrow}^\dagger a_{p\downarrow} \right), \quad (133)$$

such that  $p$  indexes the orbitals local to metal center  $A$ .

We used a well-known energy extrapolation scheme to estimate the energy errors of these MPSs [28, 47–49]. Figure 16 shows the extrapolations for the three energies using data obtained with the reverse-schedule DMRG. In the zero discarded weight limit, which represents the exact MPS, the energies are expected to be 86, 109, and 92



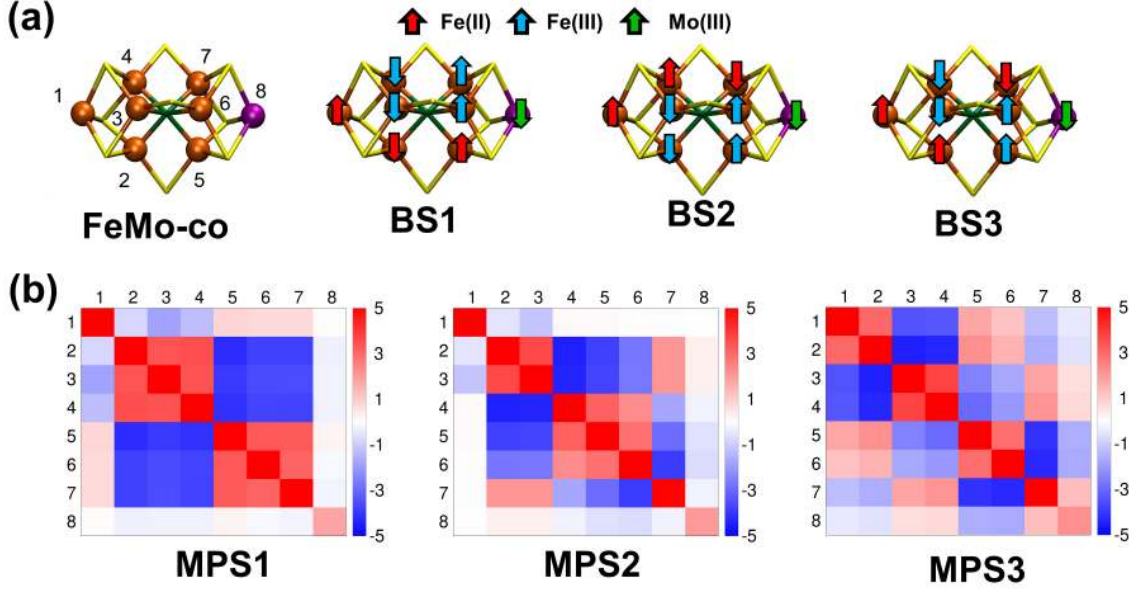


FIG. 15. (a) A schematic representation of the metal site numbering in FeMoco, along with the spin-projected broken-symmetry determinants used in the spin-projected MPS calculation. Out of 35 broken-symmetry determinants, the three with the lowest DMRG energies at a bond dimension of 2000 are represented. (b) Spin correlation matrices for the three chosen MPSs. MPS1, MPS2 and MPS3 start from the BS1, BS2, and BS3 guesses, respectively, and converged to different states.

milliHartree lower than the DMRG energies of MPS1 ( $M = 5500$ ), MPS2 ( $M = 3500$ ), and MPS3 ( $M = 3500$ ), respectively. Because these three states contain qualitatively different correlations, we consider these estimates to correspond to different low-energy eigenstates in the system. In this section we use  $M$  for the bond dimension for consistency with the literature on DMRG, in contrast to the notation  $\chi$  used for MPS preparation.

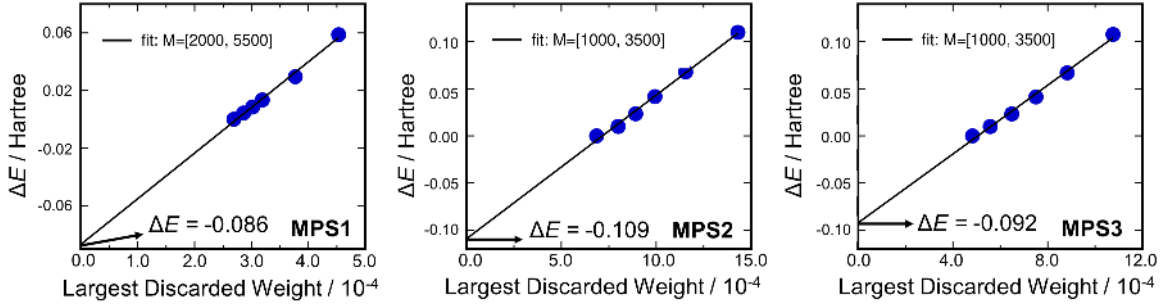


FIG. 16. Extrapolated DMRG energy for FeMoco with MPS1, MPS2, and MPS3 with respect to discarded weight, assuming a linear relationship between discarded weight and energy. The energy differences ( $\Delta E$ ) in Hartree are represented relative to the DMRG energy at the highest bond dimension used for each extrapolation, which are  $-22140.270$ ,  $-22140.249$ , and  $-22140.223$  for MPS1 ( $M = 5500$ ), MPS2 ( $M = 3500$ ), and MPS3 ( $M = 3500$ ), respectively.

### B. Overlap extrapolation protocol

Here, we present an extrapolation scheme to predict the squared overlap between a spin-adapted DMRG state with a given bond dimension  $M$  and the exact wave function,  $|\langle \Phi(M) | \Phi(\infty) \rangle|^2$ . For the extrapolation, we utilized the following two empirical linear relations:

$$\log \left( 1 - |\langle \Phi(M') | \Phi(\infty) \rangle|^2 \right) \text{ vs. } (\log(M'))^2 \quad (134)$$

$$\log \left( |\langle \Phi(M') | \Phi(M'') \rangle|^2 - |\langle \Phi(M') | \Phi(\infty) \rangle|^2 \right) \text{ vs. } (\log(M''))^2 \text{ where } M' \ll M'' \quad (135)$$

In Appendix G, we show that these linear relations are satisfied in the  $\text{Fe}_2\text{S}_2$  system where the exact wave function for the complete active space model of CAS(30e,20o) is accessible. Building on these empirical linear relations, here we demonstrate the extrapolation scheme in detail. Specifically, we show an example of estimating  $|\langle \Phi(M = 1000) | \Phi(\infty) \rangle|^2$  for the  $2\text{Fe(II)}2\text{Fe(III)}$  system. The main objective of this scheme is to accurately determine the squared overlap using data from MPSs with bond dimensions less than  $M = 1000$ . To achieve this, we first generated MPSs with bond dimensions of 800, 600, 60, 40, and 20 using a reverse sweep DMRG calculation. We then estimated  $|\langle \Phi(M' = 20) | \Phi(\infty) \rangle|^2$  using the values of  $|\langle \Phi(M' = 20) | \Phi(M'') \rangle|^2$  for  $M'' = 600, 800, 1000$  based on the second empirical linear relation of Eq. (135). We perform a linear fit of

$$\log \left( |\langle \Phi(20) | \Phi(M'') \rangle|^2 - |\langle \Phi(20) | \Phi(\infty) \rangle|^2 \right) \text{ vs. } (\log(M''))^2 \quad (136)$$

to determine the value of  $|\langle \Phi(20) | \Phi(\infty) \rangle|^2$ . In a similar fashion, we can estimate  $|\langle \Phi(M') | \Phi(\infty) \rangle|^2$  for the other bond dimensions of  $M' = 40$  and  $60$ . The empty black triangles in Fig. 17(a) represent the infidelities for these estimated values using  $M' = 20, 40, 60$ . Each empty black triangle was estimated from the blue, yellow, and green triangles directly below it. In Appendix G, we discuss the validity of this extrapolation in detail.

Finally, these estimated values were used to obtain a linear fit represented by the dotted line in Fig. 17(a), based on the first empirical linear relationship of Eq. (134). Using this line, we estimate the value of  $|\langle \Phi(1000) | \Phi(\infty) \rangle|^2$  at the black square.

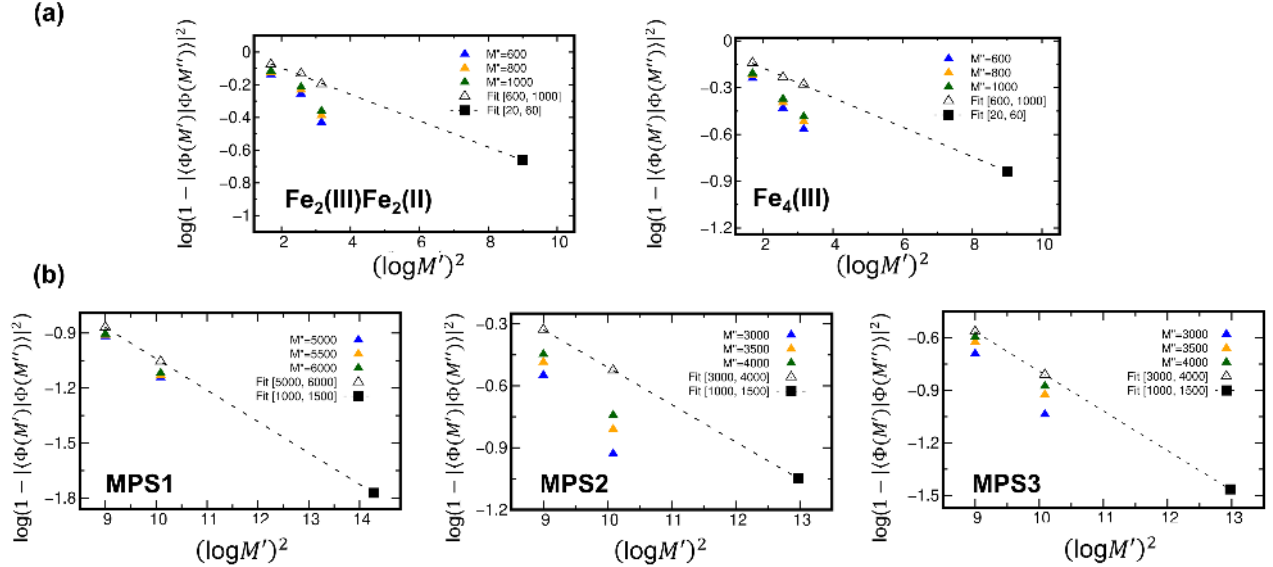


FIG. 17. Extrapolated overlap for (a)  $2\text{Fe(II)}2\text{Fe(III)}$ ,  $4\text{Fe(III)}$ , and (b)  $\text{FeMoco}$ .

We used the same extrapolation scheme to predict the overlap for the  $4\text{Fe(III)}$  and  $\text{FeMoco}$  systems. For  $4\text{Fe(III)}$ , we utilized MPSs with  $M' = 20, 40, 60$  and  $M'' = 600, 800, 1000$ . For  $\text{FeMoco}$ , we used MPSs with  $M' = 1000, 1500$  and  $M'' = 5000, 5500, 6000$  for MPS1 and  $M'' = 3000, 3500, 4000$  for MPS2 and MPS3. It is noted that we used MPS1 with  $M = 6000$  for the overlap estimate although MPS1 with  $M = 7000$  was obtained from the DMRG calculation. This was because the site of the renormalized wavefunction with  $M = 7000$  differed from the MPSs obtained through the reversed sweep, making it difficult to compute the overlap. The predicted overlap values are summarized in Table I along with the associated state preparation costs using the unitary synthesis techniques described earlier.

System	Estimated $ \langle \text{MPS}   \psi_0 \rangle $	Bond Dimension	Spatial Orbitals	MPS Toffolis	qubits
Fe <sub>2</sub> (III)Fe <sub>2</sub> (II)	0.88	1000	36	42 200 000	359
Fe <sub>4</sub> (III)	0.92	1000	36	42 200 000	359
FeMoco [43] MPS1	0.99	6000	76	1 360 000 000	833
FeMoco [43] MPS2	0.95	4000	76	733 000 000	682
FeMoco [43] MPS3	0.98	4000	76	733 000 000	682

TABLE I. Extrapolated MPS overlaps as the output of the protocol described in Section III. The physical dimension of each subsystem for the MPS is taken to be  $d = 4$  corresponding to  $\{|\emptyset\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}$ . In the case of FeMoco, the overlap is with respect to a low-energy eigenstate, which may not be the ground state. Toffoli and qubit counts are calculated using Qualtran [50].

### C. MPS preparation costs versus phase estimation costs

With the costs of performing MPS preparation, extrapolated overlap values, and number of samples using phase estimation, we are in a position to perform resource estimation for the full task of ground state energy estimation. The cost of performing block encodings for two [4Fe-4S] systems and FeMoco are determined using the tensor hypercontraction factorization of the two-electron integral tensor [27] and double factorization [51] using number operator symmetry shifting described in Ref. [52] on the one-body and two-body components of the Hamiltonian. For THC resource estimates we apply symmetry shifting only to the one-body component. We provide symmetry shifted and non-symmetry shifted block encoding costs along with high-spin coupled cluster correlation energies (obtained using PySCF [53, 54]) for a variety of cutoff parameters in Appendix F. For resource estimates we use the 1 milliHartree threshold in correlation energy differences (with respect to the non-truncated integrals) previously used in Ref. [27]. We find that for all iron-sulfur clusters studied the 1-norm value is approximately halved by symmetry shifting.

For the cost of sampling to obtain the confidence intervals, one could use Eq. (1), but in this high-overlap regime that expression tends to overestimate the cost. For a more accurate estimate one can perform the following procedure.

1. For a given number of samples  $n$ , solve Eq. (2) for  $\delta$ .
2. Given that  $\delta$ , solve Eq. (62) for  $c$ .
3. Take the cost as  $nc\lambda/\epsilon$ .

This procedure is used for a number of values of  $n$  to find the one that gives the minimum cost. Here we are using the prolate spheroidal window, which we find to give the best performance in the high-overlap regime. In the results below we also use the method in Section VC to obtain a more accurate estimate accounting for the excited states. In this high overlap regime it only gives a small correction.

Given these sampling costs, we then use this overhead multiplied by the block encoding cost. We then add  $n$  times the MPS state preparation cost to obtain the total complexity. For consistency with previous studies we select  $\epsilon = 1.0 \times 10^{-3}$  Hartree, though here we are taking this to be the confidence interval half-width, rather than the root-mean square (RMS) error as in previous studies. As a result, the accuracy requirement is somewhat more demanding, as confidence intervals are significantly wider than the RMS error.

In our estimates, we reduced the value of  $\lambda_{DF}$  with a symmetry shift of  $(\alpha_2/2)\hat{N}\hat{N} + (\alpha_1 - \alpha_2/2)\hat{N}$  where  $\hat{N}$  is the number operator and  $\alpha_1$  and  $\alpha_2$  are optimized to minimize  $\lambda_{DF}$ . With this shift, the 1-norm of the effective electron repulsion integral (ERI) tensor depends only on  $\alpha_2$ , which is optimized first subject to the constraint that the ERI tensor remains positive semi-definite. After  $\alpha_2$  has been determined,  $\alpha_1$  is chosen to minimize the 1-norm of the effective 1-electron part of the Hamiltonian which depends on both  $\alpha_1$  and  $\alpha_2$ . For THC, the LCU 1-norm  $\lambda$  is computed with a number operator symmetry shift computed as the median of  $\{f_i\}$  where  $f_i$  are eigenvalues of the one-body operator being block encoded.

Using these results for  $\lambda$ -values, we analyze the cost of Fe-S cluster ground state energy estimation with 95% and 99% confidence levels. The total Toffoli and qubit requirements are listed in Table II. Due to the overlaps being so high, few QPE samples are needed, only 2 in the example of FeMoco. We should note that this corresponds to only one candidate MPS initial guess corresponding to a specific structure of spin-coupling. Including all the BS-DFT derived MPS initial guesses considered in this work would multiply this cost by 35. This number may be reduced by pre-filtering some of the MPS initial guesses by the DMRG extrapolated energies.

System	Method	$\lambda$	BE Toffolis	Qubits	QPE Cost 95%	QPE Cost 99%
Fe <sub>2</sub> (III)Fe <sub>2</sub> (II)	THC	168.7143	9120	1149	$1.33 \times 10^{10}$	$2.45 \times 10^{10}$
	DF	154.7362	15545	3111	$2.08 \times 10^{10}$	$3.82 \times 10^{10}$
Fe <sub>4</sub> (III)	THC	164.1287	8573	1081	$8.37 \times 10^9$	$1.67 \times 10^{10}$
	DF	150.2923	15602	3113	$1.39 \times 10^{10}$	$2.77 \times 10^{10}$
FeMoco [43]	THC	781.8172	16923	2194	$7.27 \times 10^{10}$	$1.38 \times 10^{11}$
	DF	582.4211	35006	6402	$1.11 \times 10^{11}$	$2.11 \times 10^{11}$

TABLE II. The combined QPE and MPS preparation costs for confidence intervals of half-width  $\epsilon = 1.0 \times 10^{-3}$  Hartree at 95% and 99% confidence levels. The QPE costs for FeMoco are calculated using only the smallest 0.95 overlap for MPS2.

For comparison to the costings given in Ref. [27], first note that we are using the Li Hamiltonian as from Ref. [43], whereas Ref. [27] also considered the Reiher Hamiltonian from Ref. [55]. The Li Hamiltonian has higher cost to simulate, and it is those costs in Ref. [27] we should compare to. The relevant costs to compare to are therefore the right column in Table III of Ref. [27]. For THC, the Toffoli cost is  $3.2 \times 10^{10}$ , so the estimated cost here for a 95% interval is about 2.3 times higher. This factor comes from three main considerations.

1. Here we have used symmetry shifting to reduce  $\lambda$  to 781.8, which *reduces* the complexity by a factor of about 1.54 as compared to the  $\lambda$  in Ref. [27] of 1201.5. For the value of  $\lambda$ , see the line for  $M = 450$  in Table V of Ref. [27].
2. Requiring two samples doubles the complexity.
3. Requiring the measurements to provide confidence intervals of half-width  $\epsilon$  rather than RMS error of  $\epsilon$  is more demanding, and also increases the cost.

For DF, the estimated cost from Ref. [27] is  $6.4 \times 10^{10}$ , so the estimated cost for 95% confidence intervals here is about 1.7 times larger. The reduction in  $\lambda$  due to symmetry shifting for DF is more significant than for THC. The value of  $\lambda$  is given in the line for  $L = 394$  in Table XIV of Ref. [27] as 1171.2, so the improvement in  $\lambda$  here is more than a factor of 2. The other considerations for the sampling cost for DF are the same as for THC.

## VIII. CONCLUSIONS

In this work we develop improved methods for ground state energy estimation both via improved initial state preparation, and improved filtering of the initial state. For initial state preparation we developed an improved method for preparing matrix product states. This preparation is based on our new technique for synthesising general unitary operations, which improves the Toffoli count by about a factor of 7 over prior work. We then use that to construct an iterative procedure to prepare matrix product states with a substantially reduced complexity. The method for synthesising general unitary operations is of independent interest, as this is a very common task in quantum computing. Moreover, we have found that it may be possible to further improve the complexity by a procedure interspersing phasing with Hadamard gates. The drawback to that approach is that we do not have an efficient procedure to determine the phases required, so the unitaries are restricted to lower dimension (no more than about 8 qubits). In future work it may be possible to develop more efficient procedures to solve for the phases, making that a more viable approach.

For improved filtering, we have proposed two approaches for ground state energy estimation, both using quantum phase estimation whose efficiency is boosted through window functions. These window functions are chosen to minimize the error in a confidence interval, as opposed to RMS error which is more commonly considered. This choice for the phase measurements is useful as we need to perform multiple phase measurements, all of which need to avoid large error. The two methods are direct sampling and a binary search using amplitude estimation. For both we provided both asymptotic expressions and numerical estimates of the complexity. The advantage of the binary search with amplitude estimation is that it provides a square-root speedup in the overlap, though it has a larger constant factor than the direct sampling approach. This means that the direct sampling approach is preferable for the case where the initial guess has large overlap with the exact ground state, whereas the binary search approach is more advantageous in the small overlap situation. The asymptotic expressions suggest that the crossover is at  $p \sim 0.003$ , and that prediction is borne out by the numerics.

Building on the efficient MPS state preparation results along with the optimal window function analysis, we analyzed the cost of refining energy estimates using QPE initialized with an MPS wavefunction. In order to determine total costs for energy refinement in the high confidence regime we determined the overlap through an extrapolation. The extrapolation protocol uses two MPS wavefunctions to determine the infinite bond dimension overlap of a finite bond

dimension wavefunction. The extrapolation is empirical but is supported in this set of systems by verifying against true overlaps computed in the smaller FeS cluster where exact ground states can be found through large bond-dimension MPS calculations. In the case of FeMoco where the ground state energy manifold has many competing spin configurations we estimate the overlap for different MPS initial state wavefunctions that are candidates for different low-energy eigenstates. The role of QPE in this setting is then to refine the energy ordering of the states, enabling the determination of the ground-state energy. Ultimately, due to the high extrapolated overlap, achieving 95% or 99% confidence intervals only requires two samples from QPE. Improvements to block encoding LCU 1-norm through symmetry shifting results in total complexities that are only 2.3 times those of naive QPE assuming perfect overlap (and with the less demanding requirement of RMS error  $\epsilon$ ) [27].

The extrapolations and overlaps estimated here provide a concrete numerical example of a complex chemical problem where classical precomputation can be used to prepare initial states of high overlap, enabling efficient QPE. As discussed previously [12, 13], the degree of quantum advantage can then be evaluated from the relative cost of classical and quantum refinement from such an initial state. While the MPS wavefunctions considered here are attractive candidates for strongly correlated molecules up to a given finite size, other types of ansatz and techniques may also be used, particularly in the study of even larger strongly correlated molecular problems. In conclusion, our work finds that algorithms based on classical state preparation and QPE provide a practical approach in real-world examples of challenging molecular chemistry.

### ACKNOWLEDGEMENTS

DWB worked on this project under a sponsored research agreement with Google Quantum AI. DWB is also supported by Australian Research Council Discovery Projects DP210101367 and DP220101602. Y.T., L.L., and G.K.C. were supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator.

- 
- [1] Y. Xiong, S. X. Ng, G.-L. Long, and L. Hanzo, *IEEE Signal Processing Letters* **29**, 1222 (2022).
  - [2] S. Greenaway, W. Pol, and S. Sim, [arXiv:2404.01396](#) (2024).
  - [3] N. M. Tubman, C. Mejuto-Zaera, J. M. Epstein, D. Hait, D. S. Levine, W. Huggins, Z. Jiang, J. R. McClean, R. Babbush, M. Head-Gordon, and K. B. Whaley, [arXiv:1809.05523](#) (2018).
  - [4] P. J. Ollitrault, C. L. Cortes, J. F. Gonthier, R. M. Parrish, D. Rocca, G.-L. Anselmetti, M. Degroote, N. Moll, R. Santagati, and M. Streif, [arXiv:2404.08565](#) (2024).
  - [5] M. Erakovic, F. Witteveen, D. Harley, J. Günther, M. Bensberg, O. R. Meitei, M. Cho, T. Van Voorhis, M. Reiher, and M. Christandl, [arXiv:2408.01940](#) (2024).
  - [6] S. Fomichev, K. Hejazi, M. S. Zini, M. Kiser, J. F. Morales, P. A. M. Casares, A. Delgado, J. Huh, A.-C. Voigt, J. E. Mueller, and J. M. Arrazola, [arXiv:2310.18410](#) (2023).
  - [7] R. Haghshenas, J. Gray, A. C. Potter, and G. K.-L. Chan, *Physical Review X* **12**, 011047 (2022).
  - [8] G. H. Low, V. Kliuchnikov, and L. Schaeffer, *Quantum* **8**, 1375 (2024).
  - [9] S.-J. Ran, *Physical Review A* **101**, 032310 (2020).
  - [10] D. Malz, G. Styliaris, Z.-Y. Wei, and J. I. Cirac, *Physical Review Letters* **132**, 040404 (2024).
  - [11] C. Schön, E. Solano, F. Verstraete, J. I. Cirac, and M. M. Wolf, *Physical Review Letters* **95**, 110503 (2005).
  - [12] S. Lee, J. Lee, H. Zhai, Y. Tong, A. M. Dalzell, A. Kumar, P. Helms, J. Gray, Z.-H. Cui, W. Liu, M. Kastoryano, R. Babbush, J. Preskill, D. R. Reichman, E. T. Campbell, E. F. Valeev, L. Lin, and G. K.-L. Chan, *Nature Communications* **14**, 1952 (2023).
  - [13] G. K. Chan, [arXiv:2407.11235](#) (2024).
  - [14] L. Lin and Y. Tong, *PRX Quantum* **3**, 010318 (2022).
  - [15] R. Zhang, G. Wang, and P. Johnson, *Quantum* **6**, 761 (2022).
  - [16] K. Wan, M. Berta, and E. T. Campbell, *Physical Review Letters* **129**, 030503 (2022).
  - [17] Y. Dong, L. Lin, and Y. Tong, *PRX Quantum* **3**, 040305 (2022).
  - [18] G. Wang, D. S. França, R. Zhang, S. Zhu, and P. D. Johnson, *Quantum* **7**, 1167 (2023).
  - [19] Z. Ding and L. Lin, *Quantum* **7**, 1136 (2023).
  - [20] H. Ni, H. Li, and L. Ying, *Quantum* **7**, 1165 (2023).
  - [21] H. Li, H. Ni, and L. Ying, *Physical Review A* **108**, 062408 (2023).
  - [22] D. Slepian, *Journal of Mathematics and Physics* **44**, 99–140 (1965).
  - [23] H. Imai and M. Hayashi, *New Journal of Physics* **11**, 043034 (2009).
  - [24] D. Patel, S. J. S. Tan, Y. Subasi, and A. T. Sornborger, [arXiv:2403.18927](#) (2024).
  - [25] D. W. Berry, Y. Su, C. Gyurik, R. King, J. Basso, A. D. T. Barba, A. Rajput, N. Wiebe, V. Dunjko, and R. Babbush, *PRX Quantum* **5**, 010319 (2024).



- [26] L. Lin and Y. Tong, [Quantum](#) **4**, 372 (2020).
- [27] J. Lee, D. W. Berry, C. Gidney, W. J. Huggins, J. R. McClean, N. Wiebe, and R. Babbush, [PRX Quantum](#) **2**, 030305 (2021).
- [28] Z. Li, S. Guo, Q. Sun, and G. K.-L. Chan, [Nature Chemistry](#) **11**, 1026 (2019).
- [29] Z. Li and G. K.-L. Chan, [Journal of Chemical Theory and Computation](#) **13**, 2681 (2017).
- [30] D. W. Berry, M. Kieferová, A. Scherer, Y. R. Sanders, G. H. Low, N. Wiebe, C. Gidney, and R. Babbush, [npj Quantum Information](#) **4**, 22 (2018).
- [31] D. Poulin, A. Kitaev, D. S. Steiger, M. B. Hastings, and M. Troyer, [Physical Review Letters](#) **121**, 010501 (2018).
- [32] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven, [Physical Review X](#) **8**, 041015 (2018).
- [33] A. Luis and J. Peřina, [Physical Review A](#) **54**, 4564 (1996).
- [34] J. Kaiser and R. Schafer, [IEEE Transactions on Acoustics, Speech, and Signal Processing](#) **28**, 105 (1980).
- [35] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
- [36] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, [Optica](#) **3**, 1460 (2016).
- [37] V. L. Pastor, J. Lundeen, and F. Marquardt, [Optics Express](#) **29**, 38441 (2021).
- [38] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, [Physical Review Letters](#) **73**, 58 (1994).
- [39] D. W. Berry, C. Gidney, M. Motta, J. R. McClean, and R. Babbush, [Quantum](#) **3**, 208 (2019).
- [40] Y. R. Sanders, D. W. Berry, P. C. Costa, L. W. Tessler, N. Wiebe, C. Gidney, H. Neven, and R. Babbush, [PRX Quantum](#) **1**, 020312 (2020).
- [41] C. Gidney, [Quantum](#) **2**, 74 (2018).
- [42] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, [Contemporary Mathematics](#) **305**, 53 (2002).
- [43] Z. Li, J. Li, N. S. Dattani, C. Umrigar, and G. K. Chan, [The Journal of Chemical Physics](#) **150**, 024302 (2019).
- [44] H. Zhai, H. R. Larsson, S. Lee, Z.-H. Cui, T. Zhu, C. Sun, L. Peng, R. Peng, K. Liao, J. Tölle, J. Yang, S. Li, and G. K.-L. Chan, [The Journal of Chemical Physics](#) **159**, 234801 (2023).
- [45] H. Zhai and G. K.-L. Chan, [The Journal of Chemical Physics](#) **154**, 224116 (2021).
- [46] L. Cao and U. Ryde, [International Journal of Quantum Chemistry](#) **118**, e25627 (2018).
- [47] Ö. Legeza and G. Fáth, [Physical Review B](#) **53**, 14349 (1996).
- [48] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K. Chan, [The Journal of Chemical Physics](#) **142**, 034102 (2015).
- [49] C. Xiang, W. Jia, W.-H. Fang, and Z. Li, [Journal of Chemical Theory and Computation](#) **20**, 775 (2024).
- [50] M. P. Harrigan, T. Khattar, C. Yuan, A. Peduri, N. Yosri, F. D. Malone, R. Babbush, and N. C. Rubin, [arXiv:2409.04643](#) (2024).
- [51] V. von Burg, G. H. Low, T. Häner, D. S. Steiger, M. Reiher, M. Roetteler, and M. Troyer, [Phys. Rev. Res.](#) **3**, 033055 (2021).
- [52] I. Loaiza, A. M. Khah, N. Wiebe, and A. F. Izmaylov, [Quantum Science and Technology](#) **8**, 035019 (2023).
- [53] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan, [The Journal of Chemical Physics](#) **153**, 024109 (2020).
- [54] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, [Wiley Interdisciplinary Reviews: Computational Molecular Science](#) **8**, e1340 (2018).
- [55] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, [Proceedings of the National Academy of Sciences](#) **114**, 7555 (2017).

## Appendix A: Further details of higher-order approximations

Here we give further details of the asymptotic expansions used to give approximations for the Kaiser window from Subsection IV B. First, continuing the integration by parts from Eq. (49) gives

$$\begin{aligned}
 & - \int_{\pi}^{\infty} \frac{\cos 2x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx \\
 &= - \left[ \frac{\text{Ci}(2x)}{\sqrt{x^2 + \pi^2\alpha^2}} \right]_{\pi}^{\infty} - \int_{\pi}^{\infty} \frac{x \text{Ci}(2x)}{(x^2 + \pi^2\alpha^2)^{3/2}} dx \\
 &= \frac{\text{Ci}(2\pi)}{\pi\sqrt{1 + \alpha^2}} - \int_{\pi}^{\infty} \frac{x \text{Ci}(2x)}{(x^2 + \pi^2\alpha^2)^{3/2}} dx \\
 &= \frac{\text{Ci}(2\pi)}{\pi\sqrt{1 + \alpha^2}} + \left[ \frac{\cos(2x) + 2x \sin(2x) - 4x^2 \text{Ci}(2x)}{8(x^2 + \pi^2\alpha^2)^{3/2}} \right]_{\pi}^{\infty} + \int_{\pi}^{\infty} \frac{3x[\cos(2x) + 2x \sin(2x) - 4x^2 \text{Ci}(2x)]}{8(x^2 + \pi^2\alpha^2)^{5/2}} dx
 \end{aligned}$$

$$\begin{aligned}
&= \frac{\text{Ci}(2\pi)}{\pi\sqrt{1+\alpha^2}} + \frac{4\pi^2 \text{Ci}(2\pi) - 1}{8\pi^3(1+\alpha^2)^{3/2}} + \int_{\pi}^{\infty} \frac{3x[\cos(2x) + 2x \sin(2x) - 4x^2 \text{Ci}(2x)]}{8(x^2 + \pi^2\alpha^2)^{5/2}} dx \\
&= \frac{\text{Ci}(2\pi)}{\pi\sqrt{1+\alpha^2}} + \frac{4\pi^2 \text{Ci}(2\pi) - 1}{8\pi^3(1+\alpha^2)^{3/2}} - \left[ \frac{3[(2x^2 - 3) \cos(2x) - 2x(2x^2 + 3) \sin(2x) + 8x^4 \text{Ci}(2x)]}{64(x^2 + \pi^2\alpha^2)^{5/2}} \right]_{\pi}^{\infty} \\
&\quad - \int_{\pi}^{\infty} \frac{15x[(2x^2 - 3) \cos(2x) - 2x(2x^2 + 3) \sin(2x) + 8x^4 \text{Ci}(2x)]}{64(x^2 + \pi^2\alpha^2)^{9/2}} dx \\
&= \frac{\text{Ci}(2\pi)}{\pi\sqrt{1+\alpha^2}} + \frac{4\pi^2 \text{Ci}(2\pi) - 1}{8\pi^3(1+\alpha^2)^{3/2}} + \frac{3[(2\pi^2 - 3) + 8\pi^4 \text{Ci}(2\pi)]}{64\pi^5(1+\alpha^2)^{5/2}} \\
&\quad - \int_{\pi}^{\infty} \frac{15x[(2x^2 - 3) \cos(2x) - 2x(2x^2 + 3) \sin(2x) + 8x^4 \text{Ci}(2x)]}{64(x^2 + \pi^2\alpha^2)^{7/2}} dx. \tag{A1}
\end{aligned}$$

Then dividing by the normalisation and expanding in a series gives

$$\begin{aligned}
&4\sqrt{\alpha}e^{-2\pi\alpha} \left[ C_{\alpha} - \frac{5C_{\alpha}}{16\pi\alpha} - \frac{(64 + 79C_{\alpha} - 128\pi^2)}{2^9\pi^2\alpha^2} + \frac{5(64 - 323C_{\alpha} - 128\pi^2)}{2^{13}\pi^3\alpha^3} \right. \\
&\quad - \frac{195253C_{\alpha} + 128(497 - 994\pi^2 + 384\pi^4)}{2^{19}\pi^4\alpha^4} - \frac{7579195C_{\alpha} - 640(899 - 1798\pi^2 + 384\pi^4)}{2^{23}\pi^5\alpha^5} \\
&\quad \left. - \frac{2131315809C_{\alpha} + 64(6514017 - 13028034\pi^2 + 4733184\pi^4 - 655360\pi^6)}{2^{28}3\pi^6\alpha^6} + \mathcal{O}(\alpha^{-7}) \right], \tag{A2}
\end{aligned}$$

where we are giving further terms beyond that in Eq. (50).

Integrating over the tails for  $\Delta \neq 1$  gives

$$\begin{aligned}
2 \int_{(\pi/N)\sqrt{\Delta^2+\alpha^2}}^{\infty} \frac{\sin^2 \sqrt{N^2\phi^2 - \pi^2\alpha^2}}{(N^2\phi^2 - \pi^2\alpha^2)} d\phi &= \frac{2}{N} \int_{\pi\sqrt{\Delta^2+\alpha^2}}^{\infty} \frac{\sin^2 \sqrt{\phi^2 - \pi^2\alpha^2}}{(\phi^2 - \pi^2\alpha^2)} d\phi \\
&= \frac{2}{N} \int_{\pi\Delta}^{\infty} \frac{\sin^2 x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx \\
&= \frac{1}{N} \int_{\pi\Delta}^{\infty} \frac{1 - \cos 2x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx \\
&= \frac{\text{arcsinh}(\alpha/\Delta)}{\pi N\alpha} - \frac{1}{N} \int_{\pi\Delta}^{\infty} \frac{\cos 2x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx. \tag{A3}
\end{aligned}$$

Integration by parts gives

$$\begin{aligned}
&- \int_{\pi\Delta}^{\infty} \frac{\cos 2x}{x\sqrt{x^2 + \pi^2\alpha^2}} dx \\
&= - \left[ \frac{\text{Ci}(2x)}{\sqrt{x^2 + \pi^2\alpha^2}} \right]_{\pi\Delta}^{\infty} - \int_{\pi\Delta}^{\infty} \frac{x \text{Ci}(2x)}{(x^2 + \pi^2\alpha^2)^{3/2}} dx \\
&= \frac{\text{Ci}(2\pi\Delta)}{\pi\sqrt{\Delta^2 + \alpha^2}} - \int_{\pi\Delta}^{\infty} \frac{x \text{Ci}(2x)}{(x^2 + \pi^2\alpha^2)^{3/2}} dx \\
&= \frac{\text{Ci}(2\pi\sqrt{\Delta})}{\pi\sqrt{\Delta^2 + \alpha^2}} + \left[ \frac{\cos(2x) + 2x \sin(2x) - 4x^2 \text{Ci}(2x)}{8(x^2 + \pi^2\alpha^2)^{3/2}} \right]_{\pi\Delta}^{\infty} + \int_{\pi\Delta}^{\infty} \frac{3x[\cos(2x) + 2x \sin(2x) - 4x^2 \text{Ci}(2x)]}{8(x^2 + \pi^2\alpha^2)^{5/2}} dx \\
&= \frac{\text{Ci}(2\pi\Delta)}{\pi\sqrt{\Delta^2 + \alpha^2}} + \frac{4\pi^2\Delta^2 \text{Ci}(2\pi) - \cos(2\pi\Delta) - 2\pi\Delta \sin(2\pi\Delta)}{8\pi^3(\Delta^2 + \alpha^2)^{3/2}} + \int_{\pi\Delta}^{\infty} \frac{3x[\cos(2x) + 2x \sin(2x) - 4x^2 \text{Ci}(2x)]}{8(x^2 + \pi^2\alpha^2)^{5/2}} dx \\
&= \frac{\text{Ci}(2\pi\Delta)}{\pi\sqrt{\Delta^2 + \alpha^2}} + \frac{4\pi^2\Delta^2 \text{Ci}(2\pi) - \cos(2\pi\Delta) - 2\pi\Delta \sin(2\pi\Delta)}{8\pi^3(\Delta^2 + \alpha^2)^{3/2}} \\
&\quad - \left[ \frac{3[(2x^2 - 3) \cos(2x) - 2x(2x^2 + 3) \sin(2x) + 8x^4 \text{Ci}(2x)]}{64(x^2 + \pi^2\alpha^2)^{5/2}} \right]_{\pi\Delta}^{\infty} \\
&\quad - \int_{\pi\Delta}^{\infty} \frac{15x[(2x^2 - 3) \cos(2x) - 2x(2x^2 + 3) \sin(2x) + 8x^4 \text{Ci}(2x)]}{64(x^2 + \pi^2\alpha^2)^{9/2}} dx \\
&= \frac{\text{Ci}(2\pi\Delta)}{\pi\sqrt{\Delta^2 + \alpha^2}} + \frac{4\pi^2\Delta^2 \text{Ci}(2\pi) - \cos(2\pi\Delta) - 2\pi\Delta \sin(2\pi\Delta)}{8\pi^3(\Delta^2 + \alpha^2)^{3/2}}
\end{aligned}$$



$$\begin{aligned}
& + \frac{3[(2\pi^2\Delta^2 - 3)\cos(2\pi\Delta) - 2\pi\Delta(2\pi^2\Delta^2 + 3)\sin(2\pi\Delta) + 8\pi^4\Delta^4\text{Ci}(2\pi\Delta)]}{64\pi^5(\Delta^2 + \pi^2\alpha^2)^{5/2}} \\
& - \int_{\pi\Delta}^{\infty} \frac{15x[(2x^2 - 3)\cos(2x) - 2x(2x^2 + 3)\sin(2x) + 8x^4\text{Ci}(2x)]}{64(x^2 + \pi^2\alpha^2)^{7/2}} dx.
\end{aligned} \tag{A4}$$

Dividing by the normalisation gives the error as

$$\begin{aligned}
& 4\sqrt{\alpha}e^{-2\pi\alpha} \left[ C_{\alpha,\Delta} - \frac{5C_{\alpha,\Delta}}{16\pi\alpha} - \frac{(64\cos(2\pi\Delta) + 128\pi\Delta\sin(2\pi\Delta) + 79C_{\alpha,\Delta} - 128\pi^2\Delta^2)}{2^9\pi^2\alpha^2} \right. \\
& + \frac{5(64\cos(2\pi\Delta) + 128\pi\Delta\sin(2\pi\Delta) - 323C_{\alpha,\Delta} - 128\pi^2\Delta^2)}{2^{13}\pi^3\alpha^3} \\
& - \frac{195253C_{\alpha,\Delta} + 128[2\pi^2\Delta^2(79 + 192\pi^2\Delta^2) + (497 - 1152\pi^2\Delta^2)\cos(2\pi\Delta) + 2\pi\Delta(497 - 384\pi^2\Delta^2)\sin(2\pi\Delta)]}{2^{19}\pi^4\alpha^4} \\
& - \frac{7579195C_{\alpha,\Delta} + 640[2\pi^2\Delta^2(323 - 192\pi^2\Delta^2) - (899 - 1152\pi^2\Delta^2)\cos(2\pi\Delta) - 2\pi\Delta(899 - 384\pi^2\Delta^2)\sin(2\pi\Delta)]}{2^{23}\pi^5\alpha^5} \\
& \left. + \mathcal{O}(\alpha^{-6}) \right],
\end{aligned} \tag{A5}$$

where  $C_{\alpha,\Delta} = \ln(2\alpha/\Delta) + \text{Ci}(2\pi\Delta)$ . If we take  $c = \pi\sqrt{\Delta^2 + \alpha^2}$  and expand in a series in  $c$ , then we obtain

$$\begin{aligned}
& 4\sqrt{c/\pi}e^{-2c} \left[ C_{c,\Delta} - \frac{(5 - 16\pi^2\Delta^2)C_{c,\Delta}}{16c} \right. \\
& - \frac{[64\cos(2\pi\Delta) + 128\pi\Delta\sin(2\pi\Delta) + (79 + 288\pi^2\Delta^2 - 256\pi^4\Delta^4)C_{c,\Delta} - 128\pi^2\Delta^2]}{2^9c^2} \\
& + \frac{64(5 - 16\pi^2\Delta^2)[\cos(2\pi\Delta) + 2\pi\Delta\sin(2\pi\Delta)] - (1615 + 1904\pi^2\Delta^2 + 1280\pi^4\Delta^4 - 2^{12}\pi^6\Delta^6/3)C_{c,\Delta}}{2^{13}c^3} \\
& \left. + \frac{640\pi^2\Delta^2 - 2^{11}\pi^4\Delta^4}{2^{13}c^3} + \mathcal{O}(c^{-4}) \right],
\end{aligned} \tag{A6}$$

where  $C_{c,\Delta} = \ln(2c/\pi\Delta) + \text{Ci}(2\pi\Delta)$ .

## Appendix B: Slepian expansion

Here we explain the details of how to derive the higher-order terms to correct Eq. (4.4) of Slepian [22]. We start from Eq. (4.3) of Slepian, which is

$$\frac{1}{\lambda_n} \frac{\partial \lambda_n}{\partial c} = \frac{2}{c} [\psi_{0,n}(1)]^2. \tag{B1}$$

Now the error for the prolate spheroidal window is given by  $1 - \lambda_0$ , so we need  $n = 0$ . According to the expression below Eq. (4.3) of [22],  $[\psi_{0,0}(1)]^2 = N_{0,0}^2 k_3^2$ . Now  $k_3$  is given in Eq. (1.12) of [22] as

$$k_3 = e^{-c} c^{(l+1)/2} 2^{(3l+2)/2} \sqrt{\pi} P(c) Q(c). \tag{B2}$$

Now  $l = n - m$ , as per the expression below Eq. (1.10) of [22], and  $n \geq m \geq 0$ , so with  $n = 0$  we have  $m = l = 0$ . Therefore the expression for  $k_3$  simplifies to

$$k_3 = 2e^{-c} \sqrt{\pi c} 2^{(3l+2)/2} P(c) Q(c). \tag{B3}$$

For  $m = 0$  we have  $Q(c) = 1$ , according to the explanation below Eq. (1.15) of [22]. The function  $P(c)$  is given in terms of  $g$  coefficients in Eq. (1.14) of [22], which is

$$P(c) = \frac{1 + g_1^1/c + g_2^1/c^2 + \dots}{1 + g_1^2/c + g_2^2/c^2 + \dots}. \tag{B4}$$

The values of  $g$  are given as per Table IV on page 103 of that work, which have the simplified form for  $l = m = 0$

$$g_1^1 = -\frac{24}{2^8}$$

$$\begin{aligned}
g_2^1 &= -\frac{51840}{3! 2^{17}} \\
g_3^1 &= -\frac{857226240}{3! 6! 2^{21}} \\
g_1^2 &= \frac{64}{2^8} \\
g_2^2 &= \frac{165888}{3! \times 2^{17}} \\
g_3^2 &= \frac{902430720}{6! 2^{22}}.
\end{aligned} \tag{B5}$$

Next,  $N_{0,0}^2$  is given by Eq. (1.16) of [22] as

$$\frac{1}{N_{0,0}^2} = \sqrt{\pi/c} \left( 1 + \frac{3}{2^7 c^2} + \dots \right). \tag{B6}$$

This expression is missing the third-order term, meaning we cannot obtain that term in the final expression. If we assumed that third-order term was zero, we would obtain

$$\frac{2}{c} [\psi_{0,n}(1)]^2 = 2\sqrt{\pi c} e^{-2c} \left( 1 - \frac{11}{16c} - \frac{147}{512c^2} - \frac{3269}{8192c^3} - \dots \right). \tag{B7}$$

Integrating (from  $c$  to infinity) yields

$$4\sqrt{\pi c} e^{-2c} \left( 1 + \frac{1505}{1536c} - \frac{3269}{6144c^2} \right) - \frac{2177\pi \operatorname{erfc}(\sqrt{2c})}{192\sqrt{2}}. \tag{B8}$$

Expanding an asymptotic series for the erfc function then gives

$$4\sqrt{\pi c} e^{-2c} \left( 1 - \frac{7}{16c} - \frac{91}{2^9 c^2} + \frac{2177}{2^{13} c^3} \right). \tag{B9}$$

The third-order term here is different when a nonzero third-order term for  $N_{0,0}^2$  is used. Since the term found numerically is quite different, this indicates that the third-order term for  $N_{0,0}^2$  is nonzero.

Next we describe how to derive Eq. (61). In the discrete case the probability distribution for the phase error is

$$\frac{1}{2\pi} \sum_{n,m=-N}^N f(n)f(m)e^{i(n-m)\theta}. \tag{B10}$$

In the continuous limit, we replace  $x = n/(N + 1/2)$  and  $z = m/(N + 1/2)$ , and use  $\vartheta = (N + 1/2)\theta$ , to give the probability distribution for the error

$$\frac{1}{2\pi} \int_{-1}^1 dx \int_{-1}^1 dz f(x)f(z) e^{i(x-z)\vartheta}. \tag{B11}$$

with the convention that the continuous function  $f(x)$  is normalised over the interval  $[-1, 1]$ . Then the integral for the confidence level is

$$\frac{1}{2\pi} \int_{-c}^c d\vartheta \int_{-1}^1 dx \int_{-1}^1 dz f(x)f(z) e^{i(x-z)\vartheta} = \frac{c}{\pi} \int_{-1}^1 dx \int_{-1}^1 dz f(x)f(z) \operatorname{sinc}(c(x-z)). \tag{B12}$$

The solution for  $f(x)$  with maximum confidence level then corresponds to an eigenfunction of maximum eigenvalue, so that

$$(1 - \delta) f(x) = \frac{c}{\pi} \int_{-1}^1 dz f(z) \operatorname{sinc}(c(x-z)). \tag{B13}$$

In particular, we have

$$(1 - \delta) f(0) = \frac{c}{\pi} \int_{-1}^1 dz f(z) \operatorname{sinc}(cz). \tag{B14}$$

We can use this expression for the function  $f(z) = \operatorname{PS}_{0,0}(c, z)$  to give Eq. (61). The function  $\operatorname{PS}_{0,0}(c, z)$  is not normalised to 1, but that is unimportant because it appears on both sides of the equation so the normalisation cancels.

### Appendix C: Methods of calculating prolate spheroidal functions

The prolate spheroidal functions are given in various mathematical software as follows.

1. In Mathematica  $PS_{0,0}(c, z)$  is given as `SpheroidalPS[0, 0, c, z]`, and  $S_{0,0}^1(c, 1)$  is given as `SpheroidalS1[0, 0, c, 1]`. These are normalised according to the Meixner-Schärfke scheme.
2. Matlab gives the discrete prolate spheroidal sequences (i.e. for finite  $N$ ) via the function `dpss`.
3. In Python, `scipy.signal.windows.dpss` has similar functionality as `dpss` in Matlab. Python also provides `scipy.special` with `pro_ang1` and `pro_rad1` for the angular and radial prolate spheroidal functions, respectively. Unfortunately, `pro_rad1` only outputs `nan` in our testing, making it unusable.

Very similar results to those obtained in Mathematica are obtained using `dpss` in Matlab, which is a useful independent verification. To be more specific, consider the control state for phase estimation of the form

$$\sum_{n=-N}^N f(n) |n\rangle. \quad (C1)$$

Applying a phase shift and taking an inner product with the phase state gives

$$\frac{1}{\sqrt{2N+1}} \sum_{n=-N}^N f(n) e^{in(\phi - \hat{\phi})}. \quad (C2)$$

The inner product squared is

$$\frac{1}{2N+1} \sum_{n,m=-N}^N f(n)f(m) e^{i(n-m)\theta}, \quad (C3)$$

where we have replaced  $\phi - \hat{\phi}$  with  $\theta$ . The usual convention for phase measurements is to consider a slightly different normalisation convention with a continuous range of  $\theta$  values, so the probability distribution for the error is

$$\frac{1}{2\pi} \sum_{n,m=-N}^N f(n)f(m) e^{i(n-m)\theta}. \quad (C4)$$

Integrating  $\theta$  over  $[-\pi, \pi]$  then gives 1.

To obtain the probability in the confidence interval  $[-c/(N+1/2), c/(N+1/2)]$ , the integral is

$$\frac{1}{2\pi} \int_{-c/(N+1/2)}^{c/(N+1/2)} d\theta \sum_{n,m=-N}^N f(n)f(m) e^{i(n-m)\theta} = \frac{c}{\pi(N+1/2)} \sum_{n,m=-N}^N f(n)f(m) \text{sinc}\left(\frac{c(n-m)}{N+1/2}\right). \quad (C5)$$

Given  $f(n)$  from `dpss`, this expression can be used to determine the value of  $\delta$  for given  $N$ . The function `dpss` also gives  $1 - \delta$  as an output, which can be used instead of performing the explicit sum. The relative error for the value of  $\delta$  estimated with various values of  $N$  is shown in Fig. 18. It can be seen that the results using `dpss` are very accurate, even for moderate values of  $N$ . The results for  $c > 4\pi$  are obtained less accurately, because `dpss` gives  $1 - \delta$  to double precision accuracy, so cannot give accurate values for  $\delta$  below about  $10^{-14}$ . Nearly identical results are obtained using `scipy.signal.windows.dpss` in Python, though the function breaks down for  $N > 2^{15}$ .

Note that `dpss` is giving the error for a specific value of  $N$ , so the relative errors seen in Fig. 18 correspond to the difference between using the continuous window and the sampling of the continuous window for finite  $N$ . For phase estimates of relevance to quantum computing,  $N$  would be well above  $2^{10}$ , and even for that value the relative error is less than about 0.01%. This difference is less than the significant figures usually reported for complexities, so is reasonable to ignore in the analysis.

These results can be further checked by determining the maximum eigenvalue of a matrix. The maximisation of the confidence level in Eq. (C5) corresponds to finding the maximum eigenvalue of a matrix with entries

$$A_{n,m} = \frac{c}{\pi(N+1/2)} \text{sinc}\left(\frac{c(n-m)}{N+1/2}\right), \quad (C6)$$

for  $n, m$  in the range  $-N, \dots, N$ . The maximum eigenvalue of this matrix can be determined from the spectral norm. Numerical testing with a range of values of  $c$  and  $N$  yields results equal to those from `dpss` within numerical precision.

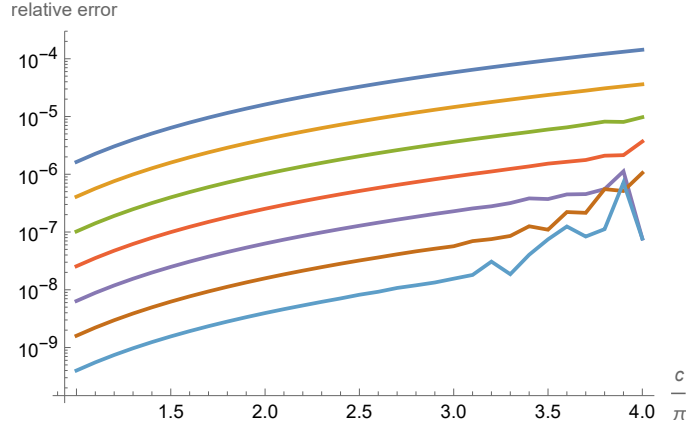


FIG. 18. The relative error for the values of  $\delta$  estimated with various values of  $N$  using `dpss` in Matlab. The lines show the results using  $N = 2^{10}$  to  $N = 2^{16}$  from top to bottom.

#### Appendix D: Combinations of excited states

To show that the probability of error should not be increased for combinations of excited states with different  $\beta$ , let us first consider the case that we just have a linear combination of the excited state at  $\beta$  around 2 and  $\beta = 0$ . That is equivalent to just considering the case where  $p$  is increased. As we are limiting the probability of error to  $q$  for *all*  $\beta$ , we have the probability of error  $\leq q$  for  $\beta = 0$ , which corresponds to  $p = 1$ . We therefore consider Eq. (87) for the error with  $p \in [p_0, 1]$  for  $p_0$  the minimum squared overlap. If there were a larger probability of error for some intermediate value of  $p$ , then we would need the first derivative of Eq. (87) to pass through zero (so there is a turning point) with the second derivative being negative (so there is a maximum).

The first derivative gives

$$\frac{dP_{\text{err}}}{dp} = n(\delta/2 - \delta_1)[p\delta/2 + (1-p)\delta_1]^{n-1} + n(\delta/2 - \delta_2)\{1 - [p\delta/2 + (1-p)\delta_2]\}^{n-1}, \quad (\text{D1})$$

so the first derivative being zero implies

$$\frac{(\delta_1 - \delta/2)}{(\delta/2 - \delta_2)} = \frac{\{1 - [p\delta/2 + (1-p)\delta_2]\}^{n-1}}{[p\delta/2 + (1-p)\delta_1]^{n-1}}. \quad (\text{D2})$$

The second derivative gives

$$\frac{d^2P_{\text{err}}}{dp^2} = n(n-1)(\delta/2 - \delta_1)^2[p\delta/2 + (1-p)\delta_1]^{n-2} - n(n-1)(\delta/2 - \delta_2)^2\{1 - [p\delta/2 + (1-p)\delta_2]\}^{n-2}, \quad (\text{D3})$$

so for the second derivative to be negative we would need

$$\frac{(\delta_1 - \delta/2)^2}{(\delta/2 - \delta_2)^2} < \frac{\{1 - [p\delta/2 + (1-p)\delta_2]\}^{n-2}}{[p\delta/2 + (1-p)\delta_1]^{n-2}}. \quad (\text{D4})$$

But, combining that with the first equality gives

$$\delta_1 - \delta/2 < \delta/2 - \delta_2. \quad (\text{D5})$$

The definitions trivially give that  $\delta_1 > \delta/2$  and  $\delta/2 < \delta_2$ , because  $\delta_1$  is the probability of the estimate being more than  $E_0 + \epsilon$  for measurements of the excited state. Moreover, for  $\beta > 2$  we have  $\delta_1 \sim 1$  and both  $\delta$  and  $\delta_2$  very small. This means that the LHS of Eq. (D5) is about 1, whereas the RHS is small, so this inequality cannot be satisfied. As this inequality was needed to obtain a larger probability of error for intermediate values of  $p$ , that is not possible.

We can perform similar reasoning for the case where there are two excited states both with  $\beta > 0$ . In that case, let us take the value of  $\delta_1$  to be a weighted linear combination of the values for the two excited states as  $\delta_1 = s\delta_1^a + (1-s)\delta_1^b$ , and similarly  $\delta_2 = s\delta_2^a + (1-s)\delta_2^b$ . Now we can perform similar reasoning for the weighting factor  $s$  as we used above for  $p$ . To have a maximum of the probability of error for  $s \in (0, 1)$  we must have the first derivative pass through zero when the second derivative is negative.

The first derivative gives

$$\frac{dP_{\text{err}}}{ds} = n(1-p)(\delta_1^a - \delta_1^b)[p\delta/2 + (1-p)\delta_1]^{n-1} + n(1-p)(\delta_2^a - \delta_2^b)\{1 - [p\delta/2 + (1-p)\delta_2]\}^{n-1}, \quad (\text{D6})$$

so the first derivative being zero implies

$$\frac{\delta_1^b - \delta_1^a}{\delta_2^a - \delta_2^b} = \frac{\{1 - [p\delta/2 + (1-p)\delta_2]\}^{n-1}}{[p\delta/2 + (1-p)\delta_1]^{n-1}}. \quad (\text{D7})$$

The second derivative gives

$$\frac{d^2 P_{\text{err}}}{ds^2} = n(n-1)(p-1)^2(\delta_1^a - \delta_1^b)^2[p\delta/2 + (1-p)\delta_1]^{n-2} - n(n-1)(p-1)^2(\delta_2^a - \delta_2^b)^2\{1 - [p\delta/2 + (1-p)\delta_2]\}^{n-2}, \quad (\text{D8})$$

so for the second derivative to be negative we would need

$$\left(\frac{\delta_1^b - \delta_1^a}{\delta_2^a - \delta_2^b}\right)^2 < \frac{\{1 - [p\delta/2 + (1-p)\delta_2]\}^{n-2}}{[p\delta/2 + (1-p)\delta_1]^{n-2}}. \quad (\text{D9})$$

Then we would require the inequality

$$\frac{\delta_1^b - \delta_1^a}{\delta_2^a - \delta_2^b} < 1. \quad (\text{D10})$$

That would imply that the RHS of Eq. (D7) is required to be less than 1, but that can be ruled out. First, note that we must have  $\delta_2 < 1 - \delta_1$ . That is because  $1 - \delta_1$  is the probability of the measurement result being *less than*  $E_0 + \epsilon$ , and  $\delta_2$  is the probability of the estimate being less than  $E_0 - \epsilon$  (which of course must be less). Moreover, we must have  $\delta < 1$ , and so we obtain

$$\begin{aligned} p\delta + (1-p)\delta_2 &< p + (1-p)(1 - \delta_1), \\ p\delta/2 + (1-p)\delta_2 &< 1 - p\delta/2 - (1-p)\delta_1, \\ 1 - [p\delta/2 + (1-p)\delta_2] &> p\delta/2 + (1-p)\delta_1, \\ \frac{1 - [p\delta/2 + (1-p)\delta_2]}{p\delta/2 + (1-p)\delta_1} &> 1, \\ \frac{\{1 - [p\delta/2 + (1-p)\delta_2]\}^{n-1}}{[p\delta/2 + (1-p)\delta_1]^{n-1}} &> 1. \end{aligned} \quad (\text{D11})$$

Thus we find that it is impossible to obtain a larger error probability by taking a combination of two excited states. Moreover, this argument shows that it is impossible to obtain a larger error probability with a combination of any number of excited states. To see that result, we can use  $\delta_1^a, \delta_2^a$  and  $\delta_1^b, \delta_2^b$  to be the probabilities resulting from disjoint sets of eigenstates. The reasoning that  $\delta_2 < 1 - \delta_1$  holds regardless of the combination of excited states, because the probability of a measurement result below  $E_0 + \epsilon$  must always be greater than the probability of a result below  $E_0 - \epsilon$ . Thus for any combination of excited states we can always split the set into two such that one of the subsets gives at least as large a value of  $P_{\text{err}}$ . Repeating this process gives a single excited state with at least as large an error probability. Thus we can obtain the maximum error probability with a single excited state, and this result upper bounds the error probability for any spectrum of excited states.

## Appendix E: Costing for LKS scheme

Here we give a detailed analysis of the cost of unitary synthesis using the method of Low, Kliuchnikov and Schaeffer [8]. For the unitary synthesis the approach of Ref. [8] is to perform a sequence of  $K$  reflections  $\mathbb{1} - 2|v_k\rangle\langle v_k|$  when  $K$  columns of the unitary need be specified (here  $K = \chi$ ), as well as a diagonal operation (phases in the computational basis). The reflections are implemented by a sequence of two state preparations (one forward and one reverse). In the following description we take the dimension to be  $N_{\text{un}}$ , which need not be a power of 2, and the number of qubits to be  $n$  so  $2^n \geq N_{\text{un}}$ . The qubits are ordered such that the most significant qubit is first. The state preparation is applied by the following procedure.

1. Perform a rotation on the first qubit to prepare  $\sqrt{p_0}|0\rangle + \sqrt{p_1}|1\rangle$ . This has Toffoli complexity  $b$  for  $b$  bits of precision for the rotation, by using a phase gradient state.
2. Use the state of the first qubit to output  $b$  bits for the rotation on the second qubit. Then use that data to perform a controlled rotation on the second qubit, then erase it. There is zero Toffoli complexity for the QROM on the single qubit, and again there is complexity  $b$  for the rotation.
3. For qubits  $k = 3$  to  $n - 1$ , use QROM on qubits 1 to  $k - 1$  to output a rotation angle. This has complexity

$$\left\lceil \frac{N_{\text{un}}}{2^{n-k+1}\Lambda_k} \right\rceil + (\Lambda_k - 1)b, \quad (\text{E1})$$

for the parameter  $\Lambda_k$  a power of 2 in the QROM. Reference [8] assumed that  $\Lambda$  is taken independent of  $k$ , but it may be adjusted to minimise complexity. Then there is complexity  $b$  for the rotation on qubit  $k$ , and erasure of the QROM needs complexity

$$\left\lceil \frac{N_{\text{un}}}{2^{n-k+1}\Lambda'_k} \right\rceil + \Lambda'_k, \quad (\text{E2})$$

for total complexity

$$\left\lceil \frac{N_{\text{un}}}{2^{n-k+1}\Lambda_k} \right\rceil + \Lambda_k b + \left\lceil \frac{N_{\text{un}}}{2^{n-k+1}\Lambda'_k} \right\rceil + \Lambda'_k, \quad (\text{E3})$$

where  $\Lambda_k$  and  $\Lambda'_k$  may be chosen independently.

4. At the end the phases are applied. This requires a QROM on all  $n$  qubits, then addition into a phase gradient register, and erasure of the QROM, with complexity

$$\left\lceil \frac{N_{\text{un}}}{\Lambda_n} \right\rceil + \Lambda_n b + \left\lceil \frac{N_{\text{un}}}{\Lambda'_n} \right\rceil + \Lambda'_n. \quad (\text{E4})$$

Steps 1 to 3 prepare a state with real amplitudes, then the final step applies phases. This final step applying phases would be applied twice between each of the  $K$  reflections. It is more efficient to combine these phases, so that there are  $K$  reflections by states with real coefficients, and  $K + 1$  diagonal phasing operations.

## Appendix F: Threshold analysis for resource estimates

### 1. $\text{Fe}_2\text{S}_2$

$n_{\text{THC}}$	$\lambda_{\text{THC}}$	$\ V - V_{\text{THC}}\ $	$E_{\text{corr}} - E_{\text{corr}}^* \text{ [mHa]}$	$C_{\text{B.E.}}$	QPE	Logical Qubits
60	104.4673	0.1431	3.7846	3903	$4.0030 \times 10^8$	575
80	105.3320	0.0400	4.9146	4222	$4.3660 \times 10^8$	580
100	105.4446	0.0219	$6.9844 \times 10^{-1}$	4414	$4.5694 \times 10^8$	581
120	105.5165	0.0124	$-1.4958 \times 10^{-1}$	4626	$4.7921 \times 10^8$	581
140	105.5543	0.0082	$-4.2948 \times 10^{-1}$	4951	$5.1306 \times 10^8$	598
160	105.5910	0.0069	$-3.8850 \times 10^{-1}$	5208	$5.3988 \times 10^8$	598
180	106.0045	0.0070	$-3.1358 \times 10^{-1}$	5418	$5.6385 \times 10^8$	1045

TABLE III. Analysis of THC rank versus accuracy of CCSD(T) for  $\text{Fe}_2\text{S}_2$  with (FeIII, FeIII) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 8$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF.

$n_{\text{THC}}$	$\lambda_{\text{THC}}$	$\ V - V_{\text{THC}}\ $	$E_{\text{corr}} - E_{\text{corr}}^*$ [mHa]	$C_{\text{B.E.}}$	QPE	Logical Qubits
60	62.7590	0.1431	$3.7846 \times 10^0$	3903	$2.4048 \times 10^{+8}$	573
80	63.6236	0.0400	$4.9146 \times 10^0$	4222	$2.6372 \times 10^{+8}$	578
100	63.7363	0.0219	$6.9769 \times 10^{-1}$	4414	$2.7620 \times 10^{+8}$	579
120	63.8082	0.0124	$-1.3363 \times 10^{-1}$	4626	$2.8979 \times 10^{+8}$	579
140	63.8460	0.0082	$-4.1996 \times 10^{-1}$	4951	$3.1033 \times 10^{+8}$	596
160	63.8827	0.0069	$-3.8494 \times 10^{-1}$	5208	$3.2663 \times 10^{+8}$	596
180	64.2961	0.0070	$-3.1175 \times 10^{-1}$	5418	$3.4200 \times 10^{+8}$	1043
200	72.3297	0.0069	$-5.5725 \times 10^{-1}$	5621	$3.9915 \times 10^{+8}$	1047

TABLE IV. Analysis of THC rank versus accuracy of CCSD(T) for  $\text{Fe}_2\text{S}_2$  with (FeII, FeIII) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 8$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF. The LCU 1-norm  $\lambda$  is computed with a number operator symmetry shift computed as the median of  $\{f_i\}$  where  $f_i$  are eigenvalues of the one-body operator being block encoded.

$\alpha_2 = 0$				$\alpha_2 = 0.1$			
DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$E_{\text{corr}} - E_{\text{corr}}^*$ [mHa]	DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$E_{\text{corr}} - E_{\text{corr}}^*$
$1.0 \times 10^{-2}$	62	0.0473	7.864	$1.0 \times 10^{-2}$	62	0.0471	8.250
$3.0 \times 10^{-3}$	80	0.0154	$8.804 \times 10^{-1}$	$1.0 \times 10^{-3}$	80	0.0159	1.606
$1.0 \times 10^{-3}$	93	0.0044	$2.496 \times 10^{-1}$	$1.0 \times 10^{-3}$	93	0.0046	$6.849 \times 10^{-1}$
$3.0 \times 10^{-4}$	111	0.0014	$1.424 \times 10^{-1}$	$1.0 \times 10^{-4}$	111	0.0014	$1.164 \times 10^{-1}$
$1.0 \times 10^{-4}$	131	0.0005	$6.639 \times 10^{-2}$	$1.0 \times 10^{-4}$	131	0.0005	$4.667 \times 10^{-2}$
$3.0 \times 10^{-5}$	152	0.0001	$6.824 \times 10^{-2}$	$1.0 \times 10^{-5}$	152	0.0001	$3.560 \times 10^{-3}$

TABLE V. Analysis of the accuracy of double factorization using the CCSD(T) correlation energy (in  $mE_h$ ) for  $\text{Fe}_2\text{S}_2$  with (FeII, FeIII) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 8$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF. Results are shown with no shift ( $\alpha_2 = 0$ ) and with a shift ( $\alpha_2 = 0.1$ ).

## 2. $\text{Fe}_4\text{S}_4$

$n_{\text{THC}}$	$\lambda_{\text{THC}}$	$\ V - V_{\text{THC}}\ $	$E_{\text{corr}} - E_{\text{corr}}^*$ [mHa]	$C_{\text{B.E.}}$	QPE	Logical Qubits
108	271.5315	1.2684	$1.9652 \times 10^1$	6888	$1.8362 \times 10^9$	937
144	292.3270	0.1294	$2.8155 \times 10^0$	7384	$2.1191 \times 10^9$	942
180	292.8878	0.0796	$1.9821 \times 10^0$	7804	$2.2440 \times 10^9$	1081
216	293.5319	0.0431	$7.5063 \times 10^{-1}$	8175	$2.3558 \times 10^9$	1083
252	293.7534	0.0301	$-1.9193 \times 10^{-1}$	8573	$2.4724 \times 10^9$	1083
288	293.8331	0.0273	$-1.7791 \times 10^{-1}$	9120	$2.6308 \times 10^9$	1151
324	293.9603	0.0217	$6.7870 \times 10^{-2}$	9590	$2.7676 \times 10^9$	1151
360	293.9936	0.0198	$2.2911 \times 10^{-1}$	10031	$2.8952 \times 10^9$	2110

TABLE VI. Analysis of THC rank versus accuracy of CCSD(T) for  $\text{Fe}_4\text{S}_4$  with (2 FeII, 2 FeIII) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 16$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF.

$n_{\text{THC}}$	$\lambda_{\text{THC}}$	$\ V - V_{\text{THC}}\ $	$E_{\text{corr}} - E_{\text{corr}}^*$ [mHa]	$C_{\text{B.E.}}$	QPE	Logical Qubits
144	166.9151	0.1294	$2.8155 \times 10^0$	7384	$1.2100 \times 10^9$	940
180	167.4759	0.0796	$1.9821 \times 10^0$	7804	$1.2831 \times 10^9$	1079
216	168.1200	0.0431	$7.5063 \times 10^{-1}$	8175	$1.3493 \times 10^9$	1081
252	168.3416	0.0301	$-1.9193 \times 10^{-1}$	8573	$1.4169 \times 10^9$	1081
288	168.4212	0.0273	$-1.7791 \times 10^{-1}$	9120	$1.5080 \times 10^9$	1149
324	168.5485	0.0217	$6.7870 \times 10^{-2}$	9590	$1.5869 \times 10^9$	1149
360	168.5817	0.0198	$2.2911 \times 10^{-1}$	10031	$1.6602 \times 10^9$	2108

TABLE VII. Analysis of THC rank versus accuracy of CCSD(T) for  $\text{Fe}_4\text{S}_4$  with (2 FeII, 2 FeIII) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 16$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF. The LCU 1-norm  $\lambda$  is computed with a number operator symmetry shift computed as the median of  $\{f_i\}$  where  $f_i$  are eigenvalues of the one-body operator being block encoded.

$\alpha_2 = 0$				$\alpha_2 = 0.1$			
DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$ E_{\text{corr}} - E_{\text{corr}}^* $ [mHa]	DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$ E_{\text{corr}} - E_{\text{corr}}^* $
$3.0 \times 10^{-3}$	152	0.0217820	1.223	$3.0 \times 10^{-3}$	152	0.0221000	1.444
$1.0 \times 10^{-3}$	181	0.0070033	$1.579 \times 10^{-1}$	$1.0 \times 10^{-3}$	181	0.0070700	$3.610 \times 10^{-1}$
$3.0 \times 10^{-4}$	222	0.0020602	$1.719 \times 10^{-1}$	$3.0 \times 10^{-4}$	222	0.0020400	$2.576 \times 10^{-1}$
$1.0 \times 10^{-4}$	260	0.0006839	$4.648 \times 10^{-2}$	$1.0 \times 10^{-4}$	260	0.0006840	$5.879 \times 10^{-2}$
$3.0 \times 10^{-5}$	312	0.0002120	$6.671 \times 10^{-3}$	$3.0 \times 10^{-5}$	312	0.0002130	$1.180 \times 10^{-2}$
$1.0 \times 10^{-5}$	365	0.0000736	$9.582 \times 10^{-3}$	$1.0 \times 10^{-5}$	365	0.0000731	$1.067 \times 10^{-2}$

TABLE VIII. Analysis of the accuracy of double factorization using the CCSD(T) correlation energy for  $\text{Fe}_4\text{S}_4$  with (2 FeII, 2 FeIII) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 16$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF. Results are shown with no shift ( $\alpha_2 = 0$ ) and with a shift ( $\alpha_2 = 0.1$ ).

$n_{\text{THC}}$	$\lambda_{\text{THC}}$	$\ V - V_{\text{THC}}\ $	$E_{\text{corr}} - E_{\text{corr}}^*$ [mHa]	$C_{\text{B.E.}}$	QPE	Logical Qubits
108	261.8652	1.1938	$3.1102 \times 10^1$	6888	$1.7708 \times 10^9$	935
144	280.7802	0.1477	$7.2471 \times 10^0$	7384	$2.0354 \times 10^9$	942
180	282.0736	0.0691	$-6.4487 \times 10^{-1}$	7804	$2.1611 \times 10^9$	1081
216	282.4604	0.0464	$-6.7771 \times 10^{-1}$	8175	$2.2670 \times 10^9$	1083
252	282.7991	0.0280	$-2.0255 \times 10^{-1}$	8573	$2.3802 \times 10^9$	1083
288	282.8973	0.0225	$-3.2773 \times 10^{-1}$	9120	$2.5329 \times 10^9$	1151
324	282.9653	0.0199	$-7.2980 \times 10^{-1}$	9590	$2.6641 \times 10^9$	1151
360	283.0203	0.0175	$-4.2608 \times 10^{-1}$	10031	$2.7872 \times 10^9$	2110

TABLE IX. Analysis of THC rank versus accuracy of CCSD(T) for  $\text{Fe}_4\text{S}_4$  with Fe-4(III) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 16$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF.

$n_{\text{THC}}$	$\lambda_{\text{THC}}$	$\ V - V_{\text{THC}}\ $	$E_{\text{corr}} - E_{\text{corr}}^*$ [mHa]	$C_{\text{B.E.}}$	QPE	Logical Qubits
108	143.1948	1.1938	$3.1102 \times 10^1$	6888	$9.6833 \times 10^8$	935
144	162.1098	0.1477	$7.2422 \times 10^0$	7384	$1.1752 \times 10^9$	940
180	163.4031	0.0691	$-6.4487 \times 10^{-1}$	7804	$1.2519 \times 10^9$	1079
216	163.7900	0.0464	$-6.8280 \times 10^{-1}$	8175	$1.3145 \times 10^9$	1081
252	164.1287	0.0280	$-2.0255 \times 10^{-1}$	8573	$1.3814 \times 10^9$	1081
288	164.2269	0.0225	$-3.2750 \times 10^{-1}$	9120	$1.4704 \times 10^9$	1149
324	164.2949	0.0199	$-7.2980 \times 10^{-1}$	9590	$1.5468 \times 10^9$	1149
360	164.3499	0.0175	$-4.2608 \times 10^{-1}$	10031	$1.6185 \times 10^9$	2108

TABLE X. Analysis of THC rank versus accuracy of CCSD(T) for  $\text{Fe}_4\text{S}_4$  with Fe-4(III) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 16$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF. The LCU 1-norm  $\lambda$  is computed with a number operator symmetry shift computed as the median of  $\{f_i\}$  where  $f_i$  are eigenvalues of the one-body operator being block encoded.

$\alpha_2 = 0$				$\alpha_2 = 0.1$			
DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$ E_{\text{corr}} - E_{\text{corr}}^* $ [mHa]	DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$ E_{\text{corr}} - E_{\text{corr}}^* $
$3.0 \times 10^{-3}$	154	0.0228617	1.920	$3.0 \times 10^{-3}$	154	0.0228752	2.908
$1.0 \times 10^{-3}$	185	0.0075771	$6.028 \times 10^{-1}$	$1.0 \times 10^{-3}$	185	0.0075597	$5.635 \times 10^{-1}$
$3.0 \times 10^{-4}$	226	0.0020788	$7.355 \times 10^{-2}$	$3.0 \times 10^{-4}$	226	0.0020692	$1.129 \times 10^{-1}$
$1.0 \times 10^{-4}$	265	0.0006872	$8.432 \times 10^{-2}$	$1.0 \times 10^{-4}$	265	0.0006910	$7.745 \times 10^{-2}$
$3.0 \times 10^{-5}$	316	0.0002163	$7.598 \times 10^{-3}$	$3.0 \times 10^{-5}$	316	0.0002147	$1.807 \times 10^{-2}$
$1.0 \times 10^{-5}$	369	0.0000738	$1.080 \times 10^{-3}$	$1.0 \times 10^{-5}$	369	0.0000740	$3.570 \times 10^{-3}$

TABLE XI. Analysis of the accuracy of double factorization using the CCSD(T) for  $\text{Fe}_4\text{S}_4$  with Fe-4(III) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 16$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF. Results are shown with no shift ( $\alpha_2 = 0$ ) and with a shift ( $\alpha_2 = 0.1$ ).



$n_{\text{THC}}$	$\lambda_{\text{THC}}$	$\ V - V_{\text{THC}}\ $	$E_{\text{corr}} - E_{\text{corr}}^*$ [mHa]	$C_{\text{B.E.}}$	QPE	Logical Qubits
108	280.7099	0.7869	$3.0641 \times 10^1$	6888	$1.8982 \times 10^9$	937
144	291.3346	0.1912	$5.4497 \times 10^0$	7384	$2.1120 \times 10^9$	942
180	293.2292	0.0799	$-3.8260 \times 10^0$	7804	$2.2466 \times 10^9$	1081
216	293.8136	0.0466	$2.2190 \times 10^0$	8175	$2.3581 \times 10^9$	1083
252	294.1204	0.0286	$1.8563 \times 10^0$	8573	$2.4755 \times 10^9$	1083
288	294.2313	0.0206	$5.1673 \times 10^{-2}$	9120	$2.6344 \times 10^9$	1151
324	294.2698	0.0192	$3.0606 \times 10^{-1}$	9590	$2.7705 \times 10^9$	1151
360	294.3255	0.0168	$-7.6609 \times 10^{-1}$	10031	$2.8985 \times 10^9$	2110

TABLE XII. Analysis of THC rank versus accuracy of CCSD(T) for  $\text{Fe}_4\text{S}_4$  with 2Fe(III)2Fe(II) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 16$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF.

$n_{\text{THC}}$	$\lambda_{\text{THC}}$	$\ V - V_{\text{THC}}\ $	$E_{\text{corr}} - E_{\text{corr}}^*$ [mHa]	$C_{\text{B.E.}}$	QPE	Logical Qubits
108	155.1929	0.7869	$3.0641 \times 10^1$	6888	$1.0495 \times 10^9$	935
144	165.8176	0.1912	$5.4497 \times 10^0$	7384	$1.2021 \times 10^9$	940
180	167.7122	0.0799	$-3.8260 \times 10^0$	7804	$1.2849 \times 10^9$	1079
216	168.2966	0.0466	$2.2190 \times 10^0$	8175	$1.3507 \times 10^9$	1081
252	168.6034	0.0286	$1.8563 \times 10^0$	8573	$1.4191 \times 10^9$	1081
288	168.7143	0.0206	$5.1673 \times 10^{-2}$	9120	$1.5106 \times 10^9$	1149
324	168.7528	0.0192	$3.0606 \times 10^{-1}$	9590	$1.5888 \times 10^9$	1149
360	168.8085	0.0168	$-7.6609 \times 10^{-1}$	10031	$1.6624 \times 10^9$	2108
396	168.8265	0.0174	$-3.3251 \times 10^{-2}$	10386	$1.7214 \times 10^9$	2110
432	168.8812	0.0082	$-3.6832 \times 10^{-2}$	10757	$1.7835 \times 10^9$	2110
468	168.9190	0.0073	$3.4664 \times 10^{-1}$	11156	$1.8501 \times 10^9$	2110
504	168.9603	0.0078	$-6.8758 \times 10^{-1}$	11561	$1.9177 \times 10^9$	2110
540	168.9765	0.0058	$-1.3886 \times 10^{-1}$	12167	$2.0184 \times 10^9$	2242

TABLE XIII. Analysis of THC rank versus accuracy of CCSD(T) for  $\text{Fe}_4\text{S}_4$  with 2Fe(III)2Fe(II) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 16$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF. The LCU 1-norm  $\lambda$  is computed with a number operator symmetry shift computed as the median of  $\{f_i\}$  where  $f_i$  are eigenvalues of the one-body operator being block encoded.

$\alpha_2 = 0$				$\alpha_2 = 0.1$			
DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$ E_{\text{corr}} - E_{\text{corr}}^* $ [mHa]	DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$ E_{\text{corr}} - E_{\text{corr}}^* $ [mHa]
$3.0 \times 10^{-3}$	152	0.0221358	$7.301 \times 10^{-1}$	$3.0 \times 10^{-3}$	152	0.0224483	$2.619 \times 10^{-1}$
$1.0 \times 10^{-3}$	183	0.0069379	$2.493 \times 10^{-1}$	$1.0 \times 10^{-3}$	181	0.0070526	$6.233 \times 10^{-1}$
$3.0 \times 10^{-4}$	222	0.0020665	$8.187 \times 10^{-2}$	$3.0 \times 10^{-4}$	222	0.0020752	$1.715 \times 10^{-2}$
$1.0 \times 10^{-4}$	260	0.0006887	$4.144 \times 10^{-2}$	$1.0 \times 10^{-4}$	260	0.0006901	$1.664 \times 10^{-2}$
$3.0 \times 10^{-5}$	312	0.0002146	$1.359 \times 10^{-2}$	$3.0 \times 10^{-5}$	312	0.0002151	$2.107 \times 10^{-2}$
$1.0 \times 10^{-5}$	365	0.0000741	$4.842 \times 10^{-3}$	$1.0 \times 10^{-5}$	365	0.0000736	$2.165 \times 10^{-2}$

TABLE XIV. Analysis of the accuracy of double factorization using the CCSD(T) for  $\text{Fe}_4\text{S}_4$  with 2Fe(III)2Fe(II) oxidation states [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 16$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF. Results are shown with no shift ( $\alpha_2 = 0$ ) and with a shift ( $\alpha_2 = 0.1$ ).

### 3. FeMoCo

$\alpha_2 = 0$				$\alpha_2 = 0.1$			
DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$ E_{\text{corr}} - E_{\text{corr}}^* $ [mHa]	DF thresh.	$n_{\text{DF}}$	$\ V - V_{\text{DF}}\ $	$ E_{\text{corr}} - E_{\text{corr}}^* $
$5.00 \times 10^{-3}$	312	0.0578095	2.414	$5.00 \times 10^{-3}$	312	1.706	0.0580414
$2.50 \times 10^{-3}$	344	0.0289546	1.470	$2.50 \times 10^{-3}$	344	1.621	0.0289732
$1.25 \times 10^{-3}$	394	0.0144793	$5.330 \times 10^{-2}$	$1.25 \times 10^{-3}$	394	$3.872 \times 10^{-1}$	0.0144793
$5.00 \times 10^{-4}$	470	0.0057589	$4.048 \times 10^{-1}$	$5.00 \times 10^{-4}$	472	$4.500 \times 10^{-1}$	0.0057756
$2.50 \times 10^{-4}$	526	0.0028081	$6.060 \times 10^{-2}$	$2.50 \times 10^{-4}$	522	$6.970 \times 10^{-2}$	0.0028139
$1.25 \times 10^{-4}$	589	0.0013772	$3.989 \times 10^{-2}$	$1.25 \times 10^{-4}$	585	$7.884 \times 10^{-3}$	0.0013765
$5.00 \times 10^{-5}$	679	0.0005509	$1.028 \times 10^{-2}$	$5.00 \times 10^{-5}$	679	$7.430 \times 10^{-3}$	0.0005514

TABLE XV. Analysis of the accuracy of double factorization using the CCSD(T) correlation energy for FeMoco [29] using a high spin  $n_{\uparrow} - n_{\downarrow} = 35$  reference for CCSD(T). CCSD(T) calculations are performed using UCCSD(T) in PySCF. Results are shown with no shift ( $\alpha_2 = 0$ ) and with a shift ( $\alpha_2 = 0.1$ ).

### Appendix G: Validation of overlap extrapolation protocol

Here we show several examples that demonstrate the validity of the overlap extrapolation protocol for the  $\text{Fe}_2\text{S}_2$  and  $\text{Fe}_4\text{S}_4$  systems. For the  $\text{Fe}_2\text{S}_2$  system, the exact wave function ( $\Phi(\infty)$ ) for the active space model of CAS(30e,20o) is accessible. We obtained the exact MPS and MPSs with several bond dimensions and calculate the overlap between them. The left and right panels of Fig. 19 show the plots corresponding to Eqs. (134) and (135), respectively. We can see that both empirical linear relations fit remarkably well.

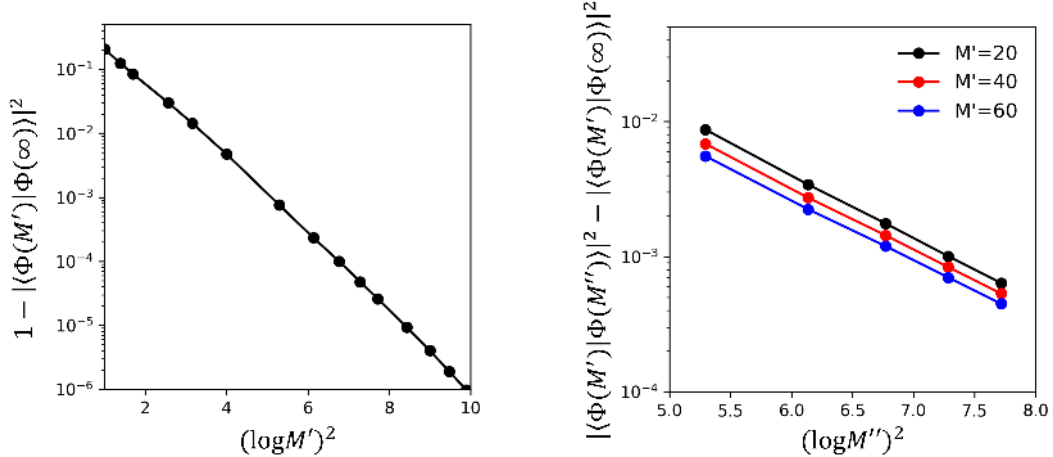


FIG. 19. Two empirical linear relations for the  $\text{Fe}_2\text{S}_2$  system. We demonstrate the first linear relation using the overlap between MPSs for  $M' = [10, 1400]$  and the exact MPS, while the second linear relation is demonstrated using the overlap between MPSs for  $M' = [20, 60]$ , MPSs for  $M'' = [200, 600]$ , and the exact MPS.

However, the exact wave functions for the complete active space models of the  $\text{Fe}_4\text{S}_4$  systems, namely CAS(54e,36o) and CAS(52e,36o), are not accessible. Therefore, while we are unable to directly confirm the empirical linear relations for these systems, Fig. 20 provides strong evidence supporting the validity of the overlap extrapolation. Each empty triangle corresponds to  $|\langle \Phi(M') | \Phi(\infty) \rangle|$  for each value of  $M' = 20, 40$ , and  $60$ , and was obtained by fitting the blue, yellow, and green triangles just below it, as discussed in Fig. 17. Based on these  $|\langle \Phi(M') | \Phi(\infty) \rangle|$  values, we applied the linear relation of Eq. (135) to predict the values of  $|\langle \Phi(M') | \Phi(M'' = 8000) \rangle|$ , corresponding to the empty circles in Fig. 20. On the other hand, we obtained the MPS for  $M'' = 8000$  using DMRG, and directly computed the overlaps between this MPS and those for  $M' = 20, 40$ , and  $60$  corresponding to the red circles. The excellent agreement between the red and empty circles provides indirect validation of  $|\langle \Phi(M') | \Phi(\infty) \rangle|$  for  $M' = 20, 40$ , and  $60$ , indicating the reliability of the extrapolation.

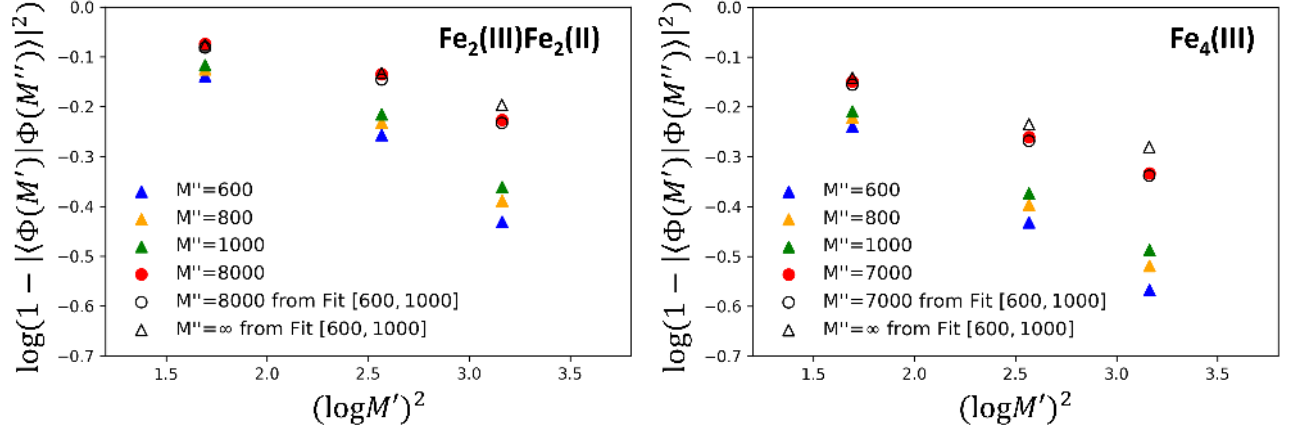


FIG. 20. Validation of the overlap extrapolation for the  $\text{Fe}_4\text{S}_4$  systems. The data from the overlaps of the MPSs for  $M' = 20, 40, 60$  and  $M'' = 600, 800, 1000, 8000$ , obtained by DMRG, are represented by the colored dots. While, the data obtained through extrapolation based on the linear relation in Eq. (135) are represented by the empty dots. The close agreement between the empty circles and red circles demonstrates the robustness of the extrapolation protocol.