

Improved Accuracy for Trotter Simulations Using Chebyshev Interpolation

Gumaro Rendon¹, Jacob Watkins², and Nathan Wiebe^{3,4}

¹Zapata Computing Inc., Boston, MA 02110, USA

²Facility for Rare Isotope Beams, Michigan State University, East Lansing, MI 48824, USA

³Department of Computer Science, University of Toronto, Toronto, ON M5S 2E4, Canada

⁴Pacific Northwest National Laboratory, Richland, WA 99352, USA

Quantum metrology allows for measuring properties of a quantum system at the optimal Heisenberg limit. However, when the relevant quantum states are prepared using digital Hamiltonian simulation, the accrued algorithmic errors will cause deviations from this fundamental limit. In this work, we show how algorithmic errors due to Trotterized time evolution can be mitigated through the use of standard polynomial interpolation techniques. Our approach is to extrapolate to zero Trotter step size, akin to zero-noise extrapolation techniques for mitigating hardware errors. We perform a rigorous error analysis of the interpolation approach for estimating eigenvalues and time-evolved expectation values, and show that the Heisenberg limit is achieved up to polylogarithmic factors in the error. Our work suggests that accuracies approaching those of state-of-the-art simulation algorithms may be achieved using Trotter and classical resources alone for a number of relevant algorithmic tasks.

1 Introduction

Quantum simulation has become, arguably, the most promising application of quantum computing in the near-term [1, 2], with the potential to provide exponential speedups for a host of problems ranging from the electronic structure [2, 3, 4, 5] to simulation of scattering dynamics within quantum field theories [6, 7, 8]. The central challenge of digital Hamiltonian simulation is, given a fixed Hamiltonian, simulation time, and error tolerance, provide a minimal-length sequence of quantum gates that approximates the unitary dynamics within that error tolerance.

Gumaro Rendon: gumaro.rendon@proton.me

Jacob Watkins: watki243@msu.edu

Nathan Wiebe: nawiebe@cs.toronto.edu

Major strides have been made in the last several years towards this goal. Several classes of methods, such as Linear Combinations of Unitaries (LCU) methods [9, 10, 11, 12, 13], qubitization [14, 15] and Trotter-based methods [16, 17, 18, 19, 20, 21] have emerged as leading approaches for simulating quantum dynamics. Unlike the other aforementioned methods, Trotter methods yield a complexity that scales with the commutators of the Hamiltonian terms [18], which can lead to substantial performance improvements for simulations of local Hamiltonian [19]. However, Trotter methods scale super-polynomially worse with the error tolerance than other existing methods [22, 11]. This makes such methods inferior to other strategies in cases where high accuracy is required.

It would be interesting if the gap in simulation accuracy between Trotter and other approaches could be bridged in a minimal way, such as classical post processing. Richardson extrapolation has been proposed as a method to mitigate algorithmic and hardware errors for simulation and linear systems [23, 24]. Recently, the approach been demonstrated on quantum hardware [25] and yielded improvements compared to Trotter simulation alone. However, there remains a gap in the theoretical understanding of the errors, though one may be able to adapt existing results from the theory of multiproduct formulas.

The present work takes a different approach to the problem, achieving improved accuracy using polynomial interpolation, hereafter referred to simply as "interpolation." In this scheme, we take Trotter simulation data at various time step sizes, then interpolate this data with a polynomial. Our estimator is then the value of this polynomial at the desired ideal of zero step size. By choosing our nodes so that the interpolation is well-conditioned, we show that a Heisenberg-limited $1/\epsilon$ scaling can be attained for estimating observables up to logarithmic factors. This is in contrast to the $1/\epsilon^{1+1/p}$ scaling obtained with a single Trotter estimate alone, where p is the order of the Trotter formula. Interestingly, our result also holds in cases where only the lowest order Trotter formula is used, providing the first poly-logarithmic scaling method with the error that uses a constant number of ancillary qubits. We study this method in two cases. The first case uses phase estimation to extract eigenvalues of the Hamiltonian; whereas the second method uses dynamical simulation and amplitude estimation to learn expectation values of observables. We further validate these methods numerically using a newly proposed unbiased phase estimation method that we call Gaussian phase estimation.

Similar to the aforementioned Richardson extrapolation, our approach is best thought of as algorithmic error mitigation. The particular error we are mitigating is imperfect Hamiltonian evolution due to the use of product formulas. Rather than being a new approach to eigenvalue or expectation value estimation, our proposed protocol relies on existing methods for, say, phase or amplitude estimation as a subroutine. We will make use of the known performance of these subroutines as we quantify the effectiveness of our full procedure.

The layout of the paper is as follows. In Section 2, we go over the basic setup of our work, including assumptions and notation. We review Suzuki-Trotter (ST) formulas and

polynomial interpolation, state an important lemma, and present the well-conditioned interpolation formulas that we use to enable our approach. Section 3 contains our analysis and applications of interpolation to eigenvalue estimation. Specifically, Section 3.1 takes a perturbative approach to deriving error bounds and complexity, and applies the results to the problem of estimating Trotter error on a quantum computer. Section 3.2 approaches the problem through complex analysis and the Bernstein ellipse, and applies the analysis to the newly proposed Gaussian phase estimation. In Section 4, we present our main results involving extrapolation of expectation values of time-evolved observables. This case is conceptually distinct from the case of phase estimation, because here the time evolutions are generally long. In Section 5 we validate our claims numerically for small instances of transverse Ising models. Finally, we conclude in Section 6 and discuss future avenues of research.

2 Setup and Notation

The algorithms presented in this paper are simply an application of standard polynomial interpolation to data obtained from Trotter simulations. These two pieces are, in many respects, disjoint, and the interpolation can be thought of as a form of classical post processing. Thus, we will begin this section with a review of the quantum part, Trotterization, then move to interpolation.

2.1 Trotter

Consider a Hamiltonian H on a collection of n qubits, decomposed in a specified way into a sum of m terms.

$$H = \sum_{j=1}^m H_j \tag{2.1}$$

Notice that both the Hamiltonian *and* decomposition are specified, as the decomposition is not unique and can have implications for algorithmic performance. Questions about optimal decompositions are entirely neglected in this work, as well as questions of how one maps a quantum system of interest (such as a molecule or crystal) onto a set of n qubits appropriately.

In our results, as a proxy for the true simulation cost, we will quote the number of exponentials $U_j = e^{-iH_j t}$ used to approximate e^{-iHt} . Of course, this only indicates the true cost when each U_j is computationally cheap. In a number of relevant situations, the Hamiltonian is k -local or sparse, and it is possible to identify decompositions such that each U_j has constant cost and m grows polynomially with respect to the number of qubits n . For such Hamiltonians, there exist efficient quantum algorithms to approximate the time evolution operator $U(t) = e^{-iHt}$ to some desired precision ϵ , many of which rely on Trotter formulas. Standard examples of Trotter formulas include

the first-order formula

$$S_1(t) := \prod_{j=1}^m e^{-iH_j t}, \quad (2.2)$$

the second-order symmetric formula

$$S_2(t) := e^{-iH_1 t/2} \dots e^{-iH_m t/2} e^{-iH_m t/2} \dots e^{-iH_1 t/2}, \quad (2.3)$$

and more generally, the order $2k$ symmetric ST formula, defined recursively as

$$S_{2k}(t) := [S_{2k-2}(u_k t)]^2 S_{2k-2}((1 - 4u_k)t) [S_{2k-2}(u_k t)]^2 \quad (2.4)$$

for every $k \in \mathbb{Z}_+ \setminus \{1\}$, where $u_k := (4 - 4^{1/(2k-1)})^{-1}$ [26]. Though many kinds of Trotter formulas exist, the ST formulas (2.4) will be our primary tool in this work for several reasons. First, they are symmetric under naive¹ time reversal

$$S_{2k}(-t) = S_{2k}(t)^\dagger \quad (2.5)$$

which will allow us to effectively double the number of interpolation points. Second, the order of the formula $2k$ can be taken arbitrarily large, and therefore our results will apply to quite a general class. Finally, the lowest order $k = 1$ symmetric formula is actually practical, having small constant factors, and therefore our bounds will have something to say about actual simulations.

Trotter formulas approximate $U(t)$ only in a neighborhood around $t = 0$. Thus, the standard trick is to divide the time interval $[0, t]$ into r subintervals, such that each interval is sufficiently small that the Trotter approximation is valid, then string together these simulations. For the simple case of a uniform mesh of r subintervals, this becomes

$$S_{2k}(t/r)^r = U(t) + O(t^{2k+1}/r^{2k}) \quad (2.6)$$

where big O is understood as taking r large. Clearly, we have that $\lim_{r \rightarrow \infty} S_{2k}(t/r)^r = U(t)$.

Rather than thinking about the number of steps r tending to infinity, it is simpler for our subsequent analysis to consider $s = 1/r$ as a "dimensionless step size," and instead think about $s \rightarrow 0$. In terms of s , we define

$$\tilde{U}_s(t) := S_{2k}(st)^{1/s}. \quad (2.7)$$

as the approximate evolution operator for $s \neq 0$. The discontinuity at $s = 0$ in (2.7) may be filled by the exact evolution $\tilde{U}_0(t) := U(t)$. Though we defined s as a reciprocal integer, definition (2.7) together with $\tilde{U}_0(t)$ suggests an extension to allow s to be real-valued. In fact, the resulting function \tilde{U}_s is smooth on a neighborhood of $s = 0$, a fact that will allow us to precisely characterize the interpolation error. For our purposes,

¹Not to be confused with actual time reversal, which is an antilinear operator.

we will only consider $|s| \leq 1$. When $1/s$ is not an integer, we may implement \tilde{U}_s using fractional queries [27] by splitting $1/s$ into integer and fractional parts.

$$1/s = r + f \tag{2.8}$$

Here, $r = \text{rnd}(1/s) \in \mathbb{Z}$ is $1/s$ rounded to the nearest integer, and $f \in [-1/2, 1/2]$. Finally, we note that \tilde{U}_s is an even function of s , which we will make use of to cut the number of interpolation points in half by reflecting across $s = 0$.

Prior work has demonstrated the value of considering the effective Trotter Hamiltonian in the analysis of Trotter formulas [28]. This approach also helps us calculate high order derivatives of \tilde{U}_s as needed for our error bounds. We thus define an effective Hamiltonian

$$\tilde{H}_s := \frac{i}{st} \log S_{2k}(st) \tag{2.9}$$

so that

$$\tilde{U}_s(t) = e^{-i\tilde{H}_s t}. \tag{2.10}$$

Note that \tilde{H}_s depends on t as well, though this dependence will be left implicit. For the purposes of bounding the interpolation error, we require a bound on the norm of \tilde{H}_s . This is supplied by the following lemma.

Lemma 1. *In the notation introduced above, let s be chosen such that*

$$k(5/3)^k m \max_{l \in [1, m]} \|H_l\| |s| t \leq \pi/20. \tag{2.11}$$

Then the following bound on the derivatives of \tilde{H}_s with respect to s holds.

$$\|\partial_s^n \tilde{H}_s\| \leq 2t^{-1} n^n (e^2 k (5/3)^k m \max_{l \in [1, m]} \|H_l\| t)^{n+1}.$$

The proof of this lemma is technical, so it is relegated to Appendix A. Note that our bounds are uniformly worse for larger k , i.e., higher order ST formulas. Assuming that this is not an artifact of our mathematical treatment, this suggests low order formulas are unconditionally preferred over high order ones for interpolation. Numerical studies could help determine the true impact of higher order formulas on the interpolation procedure.

2.2 Polynomial Interpolation

Having introduced ST formulas and stated a needed lemma, we turn our attention to interpolation. Essentially, our goal is to use interpolation to "extrapolate" to the ideal

of $s = 0$.² There are many quantities that we could be interested in extrapolating, including the eigenvalues $\lambda_i(s)$ of the effective Hamiltonian,

$$\tilde{H}_s |\lambda_i(s)\rangle = \lambda_i(s) |\lambda_i(s)\rangle \quad (2.12)$$

and expectation values of time-evolved observables.

$$\begin{aligned} \langle O_s(t) \rangle &= \text{Tr}(\rho O_s(t)) \\ O_s(t) &:= \tilde{U}_s(t)^\dagger O \tilde{U}_s(t) \end{aligned} \quad (2.13)$$

These will be the primary focus of the present work. No matter what quantity we are interested in, we'll assume it was obtained from ST on a quantum computer for the purposes of subsequent cost analysis. However, in principle our approach should work for any Trotter scheme, not just ST.

While the interpolation is classical and independent of the method in which the data is generated, we will assume a quantum simulation was used when considering the computational cost. We assume all quantum operations are executed perfectly, including the exponentials $\exp(-iH_j t)$ for simulation. Thus, the primary sources of error are the interpolation error and error in the calculation of the data points (e.g. the Hamiltonian energies or expectation values at various points s_i). Error in the data points may arise from hardware noise, but even in its absence, a measurement protocol such as phase estimation induces a systematic error that cannot be removed. In this work, we only account for the latter: algorithmic errors. This is not to say that the interpolation method could not be applied to noisy quantum systems, but rather that our cost analysis does not account for it.

We now describe the interpolation framework. Let $f \in C^\infty([-a, a])$ be a smooth, real-valued function of a single variable $s \in [-a, a]$ and suppose we have calculated f for n distinct points $s_1, s_2 \dots s_n \in [-a, a]$. That is, we have data in the form of a set of tuples $D = \{(s_i, f_i)\}_{i=1}^n$, where $f_i = f(s_i)$. Let $P_{n-1}f$ be the unique $(n-1)$ -degree polynomial interpolating D , i.e. $P_{n-1}f(s_i) = f_i$ for each $i = 1, \dots, n$. For any $s \in [-a, a]$, standard results in polynomial interpolation [29] tell us that the signed error is given by

$$E_{n-1}(s) := f(s) - P_{n-1}f(s) = \frac{f^{(n)}(\xi)}{n!} \omega_n(s) \quad (2.14)$$

for some $\xi \in I_s$, where $I_s \subset [-a, a]$ is the smallest interval containing s and the interpolation points $\{s_i\}$. Throughout this work, superscripts such as in $f^{(n)}$ will refer to n th-order derivatives. The n th degree nodal polynomial $\omega_n(s)$ is defined as the unique monic polynomial with zeros at the interpolation points.

$$\omega_n(s) := \prod_{i=1}^n (s - s_i) \quad (2.15)$$

²We occasionally interchange between the terminology "extra-" and "interpolation." We view our method as an extrapolation beyond the data using a numerical technique commonly known as polynomial interpolation.

Our estimate for $f(0)$ is $P_{n-1}f(0)$. Since we are interested in $s = 0$, ω_n becomes a (signed) product of the interpolation points. We can bound the interpolation error $E_n(0)$ in a way that is independent of the precise value of ξ (which is unknown and difficult to find) by maximizing over $\xi \in I_s$.

$$|E_{n-1}(0)| \leq \max_{s \in I_s} \frac{|f^{(n)}(s)|}{n!} \prod_{i=1}^n |s_i| \quad (2.16)$$

Much of the technical work in this paper involves finding suitable bounds on the size of the derivatives $f^{(n)}$ for different forms of f . For example, in the eigenvalue setting³ of equation (2.12), $f(s) = \lambda(s)$ while for the expectation values of equation (2.13), $f(s) = \langle O_s(t) \rangle$.

For reasons which we discuss in the next subsection, we choose the Chebyshev nodes on $[-a, a]$ as our interpolation points.

$$s_i = a \cos\left(\frac{2i-1}{2n}\pi\right) \quad (2.17)$$

This allows us to specialize our interpolation error in the manner described in the following lemma.

Lemma 2. *Let $s_i, i = 1, 2, \dots, n$ be the collection of Chebyshev interpolation points on the interval $[-a, a]$. In the notation above, we have*

$$|E_{n-1}(0)| \leq \max_{s \in [-a, a]} |f^{(n)}(s)| \left(\frac{a}{2n}\right)^n. \quad (2.18)$$

Proof. For n odd, $s = 0$ is one of the interpolation points, so the error is zero and the bound holds automatically. Hereafter, we only consider n even (which will be the case of practical interest).

Using the generic bound (2.16) with the Chebyshev nodes,

$$|E_{n-1}(0)| \leq \max_{\xi \in [-a, a]} |f^{(n)}(\xi)| \frac{1}{n!} a^n \prod_{i=1}^n \left| \cos\left(\frac{2i-1}{2n}\pi\right) \right|. \quad (2.19)$$

To obtain the lemma, we just need to appropriately bound the product of cosines. Since n is even, $n = 2m$ for some $m \in \mathbb{Z}_+$. Moreover, we have a reflectional symmetry about m , in the sense that

$$\left| \cos\left(\frac{2i-1}{2n}\pi\right) \right| = \left| \cos\left(\frac{2(n-i+1)-1}{2n}\pi\right) \right|. \quad (2.20)$$

³Smoothness assumptions can be violated by eigenvalue "level crossings." Our results will require a minimum gap as $s \rightarrow 0$ to this and other technical problems.

Hence, we only need to take the product over $i = 1, \dots, m$ and square it.

$$\prod_{i=1}^n \left| \cos \left(\frac{2i-1}{2n} \pi \right) \right| = \left(\prod_{i=1}^m \cos \left(\frac{2i-1}{4m} \pi \right) \right)^2 \quad (2.21)$$

To proceed further, let's reindex the remaining product by $i \rightarrow m - i + 1$. This gives

$$\begin{aligned} \prod_{i=1}^m \cos \left(\frac{2i-1}{4m} \pi \right) &= \prod_{i=1}^m \cos \left(\frac{\pi}{2} - \frac{2i-1}{4m} \pi \right) \\ &= \prod_{i=1}^m \sin \left(\frac{2i-1}{4m} \pi \right) \\ &\leq \prod_{i=1}^m \frac{2i-1}{4m} \end{aligned} \quad (2.22)$$

where we used the fact that $\sin(x) \leq x$ for all $x \geq 0$. Factoring out the denominator from the product, the remaining terms become a double factorial.

$$\prod_{i=1}^m \frac{2i-1}{4m} = \frac{(2m-1)!!}{(4m)^m} \quad (2.23)$$

The double factorial can be bounded as follows.

$$(2m-1)!!^2 \leq (2m-1)!!(2m)!! = 2m! \quad (2.24)$$

so that $(2m-1)!! \leq \sqrt{(2m)!}$. Returning to the original product of equation (2.21), and reintroducing $n = 2m$, the resulting bound is

$$\prod_{i=1}^n \left| \cos \left(\frac{2i-1}{2n} \pi \right) \right| \leq \left(\frac{\sqrt{n!}}{(2n)^{n/2}} \right)^2 = \frac{n!}{(2n)^n} \quad (2.25)$$

Reinserting this result into the last line of equation (2.19) gives the bound stated in the lemma. \square

Though Chebyshev interpolation enjoys nice mathematical properties, it presents a challenge for Trotter simulation because of the need for noninteger time steps in equation (2.7). In the face of this, there are several options one could take: restricting to integer time steps, or perform fractional queries using, say the Quantum Singular Value Transformation (QSVT).

First, consider restricting to integer time steps, by gathering data at the nearest reciprocal integer $1/\tilde{r}$ to the Chebyshev node s . For symmetrical interval $[-a, a]$, this distance goes as $O(a^2)$ as $a \rightarrow 0$. From here, one could either (a) take the estimate for $f(1/\tilde{r})$ as the estimate for $f(s)$, accruing some error in the process, or (b) perform the interpolation at the *approximate* Chebyshev nodes given by the collection of points $1/r_i$. Unfortunately, for our purposes, option (a) leads to unacceptable errors of order $O(a)$ in

the data, eliminating accuracy gains. As for option (b), it is possible to use robustness results on Chebyshev interpolation [30] to argue that almost-Chebyshev nodes should be almost as well-conditioned. Again, however, we find that our scaling of the number of nodes is such that the node displacements must be quite small, leading again to poor scaling.

Because of this, for most of this work we choose to invoke access to fractional queries using the QSVT [27]. A notable exception to this is in application to Gaussian phase estimation in Section 3.2.3, where the error in the fractional part can be mitigated in that context. While fractional queries increase the overhead compared to Trotter alone, this overhead is a constant.

2.3 Stability Analysis and Interpolation

Polynomial interpolation is a valuable numerical tool, but some implementations can lead to numerical instability [31]. However, the situation is not as bad as often presented in textbooks [32]. While linear algebraic approaches involving Vandermonde matrices suffer instability for high degree polynomials [33], methods such as barycentric formulas are provably stable with respect to floating point arithmetic [34].

A particularly important consideration is the choice of interpolation nodes. It is well known that equally spaced nodes can lead to the Runge phenomenon: rapid oscillations near the ends of the interval that grow with polynomial degree [29]. These oscillations can be overcome with a superior choice of nodes, such as the zeros of the Chebyshev polynomials. Interpolations done with this set of nodes are guaranteed to converge to functions that are Lipschitz continuous as $n \rightarrow \infty$. Moreover, they are well-conditioned in the sense of small errors in the data values. Finally, because they anti-cluster around $s = 0$, they are relatively cheap to compute with Trotter formulas. In this work, we will always interpolate at the n th-degree Chebyshev nodes, or approximations thereof, on a symmetric interval $[-a, a]$ about the origin, defined in (2.17). We choose even n so as to avoid the origin (which has infinite cost to compute), and also utilize the reflectional symmetry of $f(s)$.

To compute the interpolant $P_{n-1}f$ linear algebraically, we overcome the limitations of the standard Vandemonde approach by expanding in terms of orthonormal Chebyshev polynomials rather than monomials x^j .

$$P_{n-1}f(s) = \sum_{j=0}^{n-1} c_j p_j(s). \quad (2.26)$$

Here, p_j is defined by

$$p_j(s) := \begin{cases} \sqrt{\frac{1}{n}} T_0(s), & j = 0 \\ \sqrt{\frac{2}{n}} T_j(s), & j = 1, 2, \dots \end{cases} \quad (2.27)$$

where T_j is the standard j th Chebyshev polynomial.

$$T_j(x) := \cos(j \cos^{-1} x) \quad (2.28)$$

By orthonormality, we are referring to the condition [35]

$$\sum_{k=1}^n p_i(s_k)p_j(s_k) = \delta_{ij} \quad (2.29)$$

for all $0 \leq i, j < n$, with s_k being the zeros of T_n given in (2.17). This immediately implies the matrix

$$\mathbf{V} := \begin{pmatrix} p_0(s_1) & p_1(s_1) & \cdots & p_{n-1}(s_1) \\ p_0(s_2) & p_1(s_2) & \cdots & p_{n-1}(s_2) \\ \vdots & \vdots & \ddots & \vdots \\ p_0(s_n) & p_1(s_n) & \cdots & p_{n-1}(s_n) \end{pmatrix} \quad (2.30)$$

is orthogonal, and therefore has condition number $\kappa(\mathbf{V}) := \|\mathbf{V}\| \|\mathbf{V}^{-1}\|$ equal to one. This is the source of well-conditioning in our approach. The coefficients $c = (c_0, c_1, \dots, c_{n-1})$ in equation (2.26) satisfy

$$y = \mathbf{V}c \quad (2.31)$$

for the vector of values $y = (f(s_1), f(s_2), \dots, f(s_n))$, since $P_{n-1}f$ is an interpolant. Hence, $c = \mathbf{V}^T y$ gives the vector of coefficients.

We now develop our argument for well-conditioning. Unless otherwise subscripted, all logarithms are natural.

Lemma 3. *Let s_1, s_2, \dots, s_n be the standard Chebyshev nodes on $[-a, a]$ (2.17) with n even. Then the nodes satisfy*

$$\sum_{k=1}^n \frac{1}{|s_k|} \leq \frac{4n}{\pi a} (\gamma + \log(2n + 2)),$$

where $\gamma \approx 0.577$ is the Euler-Mascheroni constant.

Proof. We focus on the case $a = 1$, since the general result follows by a simple rescaling. Because sine and cosine are phase shifted by $\pi/2$,

$$\sum_{k=1}^n \frac{1}{|s_k|} = \sum_{k=1}^n \frac{1}{\left| \cos\left(\frac{2k-1}{2n}\pi\right) \right|} = \sum_{k=1}^n \frac{1}{\left| \sin\left(\frac{n-2k+1}{2n}\pi\right) \right|}. \quad (2.32)$$

Taking advantage of the symmetry about $s = 0$,

$$\sum_{k=1}^n \frac{1}{\left| \sin\left(\frac{n-2k+1}{2n}\pi\right) \right|} = 2 \sum_{k=1}^{n/2} \frac{1}{\sin\left(\frac{2k-1}{2n}\pi\right)}. \quad (2.33)$$

Next, we use the lower bound

$$\sin x \geq x/2 \quad (0 \leq x \leq \pi/2) \quad (2.34)$$

in order to bound the terms of the sum.

$$\begin{aligned} 2 \sum_{k=1}^{n/2} \frac{1}{\sin\left(\frac{2k-1}{2n}\pi\right)} &\leq \frac{8n}{\pi} \sum_{k=1}^{n/2} \frac{1}{2k-1} \\ &= \frac{8n}{\pi} \left(H_n - \frac{1}{2}H_{n/2}\right) \end{aligned} \quad (2.35)$$

Here, H_n denotes the n th harmonic number. From the relation $H_n = \gamma + \psi(n+1)$, where ψ is the digamma function,

$$H_n - \frac{1}{2}H_{n/2} = \gamma/2 + \psi(n+1) - \frac{1}{2}\psi(n/2+1). \quad (2.36)$$

Moreover, since $\psi(x) \in (\log(x - \frac{1}{2}), \log x)$ for any $x > 1/2$, this is upper bounded by

$$H_n - \frac{1}{2}H_{n/2} < \gamma/2 + \log(n+1) - \frac{1}{2}\log\left(\frac{n+1}{2}\right) = \frac{\gamma + \log 2}{2} + \frac{\log(n+1)}{2}. \quad (2.37)$$

Reinserting this into (2.35), one obtains the bound

$$\sum_{k=1}^n \frac{1}{|s_k|} \leq \frac{4n}{\pi} (\gamma + \log(2n+2)) \quad (2.38)$$

The general lemma follows from a rescaling by $1/a$. □

Remark. Observe that $1/|s_k|$ is essentially the number of Trotter steps to compute the k th interpolation point. Thus, Lemma 3 amounts to a bound on the total number of Trotter steps, and we see this grows as $O(a^{-1}n \log n)$.

Lemma 4. Let $p(s) = (p_0(s), p_1(s), \dots, p_{n-1}(s))$ be a vector of (normalized) Chebyshev polynomials on $[-a, a]$. Then,

$$\|\mathbf{V}p(0)\|_1 < \frac{2}{\pi} \log(n+1) + 1$$

where $\|\cdot\|_1$ denotes the vector 1-norm.

Proof. Let $d(s) = \mathbf{V}p(s)$. For each $k = 1, 2, \dots, n$ we have

$$\begin{aligned} d_k(s) &= \sum_{j=0}^{n-1} \mathbf{V}_{kj} p_j(s) = \sum_{j=0}^{n-1} p_j(s_k) p_j(s) \\ &= \frac{1}{n} + \frac{2}{n} \sum_{j=1}^{n-1} \cos\left(j \left(\frac{2k-1}{2n}\pi\right)\right) \cos(j \cos^{-1}(s)). \end{aligned} \quad (2.39)$$

At $s = 0$, $\cos(j \cos^{-1}(0)) = \cos(j\pi/2)$, which is zero for odd j . Hence,

$$\begin{aligned} d_k(0) &= \frac{1}{n} + \frac{2}{n} \sum_{j=2, \text{even}}^{n-2} \cos\left(j \left(\frac{2k-1}{2n}\pi\right)\right) (-1)^{j/2} \\ &= \frac{1}{n} + \frac{2}{n} \sum_{j'=1}^{n/2-1} (-1)^{j'} \cos\left(\pi j' \frac{2k-1}{n}\right). \end{aligned} \quad (2.40)$$

The sum can be evaluated exactly (the authors used Mathematica), yielding

$$d_k(0) = \frac{1}{n} - \frac{2}{n} \left(\frac{1 - \cos((k + n/2)\pi) \tan(\pi \frac{2k-1}{2n})}{2} \right) \quad (2.41)$$

$$= \frac{1}{n} - \frac{1}{n} \left(1 - (-1)^{k+n/2} \tan\left(\frac{2k-1}{2n}\pi\right) \right) \quad (2.42)$$

$$= \frac{1}{n} (-1)^{k+n/2} \tan\left(\frac{2k-1}{2n}\pi\right). \quad (2.43)$$

With coefficients in hand, we now compute the one norm of $d(0)$.

$$\|d(0)\|_1 = \frac{1}{n} \sum_{k=1}^n \left| \tan\left(\frac{2k-1}{2n}\pi\right) \right| \quad (2.44)$$

We have a reflectional symmetry about $k \rightarrow n - k + 1$, allowing us to cut the sum in half and remove the absolute value sign.

$$\begin{aligned} \|d(0)\|_1 &= \frac{2}{n} \sum_{k=1}^{n/2} \tan\left(\frac{2k-1}{2n}\pi\right) \\ &= \frac{1}{m} \sum_{k=1}^m \tan\left(\frac{2k-1}{2m}\frac{\pi}{2}\right) \end{aligned} \quad (2.45)$$

Here, $m \equiv n/2$. We observe that the sum increases as k approaches m due to the first order pole at $\pi/2$. We can upper bound $\tan(x)$, and therefore the sum above, as follows.

$$\begin{aligned} \frac{1}{m} \sum_{k=1}^m \tan\left(\frac{2k-1}{2m}\frac{\pi}{2}\right) &\leq \frac{1}{m} \sum_{k=1}^m \frac{1}{\frac{\pi}{2} - \frac{\pi}{2} \left(\frac{2k-1}{2m}\right)} = \frac{4}{\pi} \sum_{k=1}^m \frac{1}{2(m-k)+1} \\ &= \frac{4}{\pi} \sum_{j=1}^m \frac{1}{2j-1} \end{aligned} \quad (2.46)$$

In the last line, we reindexed by $j = m - k + 1$. Borrowing the reasoning from the prior lemma,

$$\frac{4}{\pi} \sum_{j=1}^m \frac{1}{2j-1} < \frac{2}{\pi} (\gamma + \log(2n+2)). \quad (2.47)$$

Tracing back, this is an upper bound on $\|d(0)\|_1$. Hence,

$$\|d(0)\|_1 < \frac{2}{\pi} \log(n+1) + \frac{2(\gamma + \log(2))}{\pi} < \frac{2}{\pi} \log(n+1) + 1 \quad (2.48)$$

□

The benefit of well-conditioning comes from relaxing the need to have exquisitely precise data to achieve good interpolations. This property is captured by the following theorem.

Theorem 5. *Let $y = (f(s_1), f(s_2), \dots, f(s_n))^T$, and let $\tilde{y} \in \mathbb{R}^n$ be an approximation of y in the sense that, for all $1 \leq j \leq n$, $|f(s_j) - \tilde{y}_j| \leq \epsilon / (\frac{2}{\pi} \log(n+1) + 1)$ with probability at least $1 - \delta/n$. Let $p(s) = (p_0(s), \dots, p_{n-1}(s))^T$ be the vector of orthonormal Chebyshev polynomials. Then $\tilde{y}^T \mathbf{V} p(s)$ is an estimate of the interpolant $P_{n-1}f(s)$ at $s = 0$ to precision*

$$\left| P_{n-1}f(0) - \tilde{y}^T \mathbf{V} p(0) \right| \leq \epsilon$$

with probability at least $1 - \delta$.

Proof. First, observe that $P_{n-1}f(s) = p(s)^T c = p(s)^T \mathbf{V}^T y$ by the discussion surrounding (2.31). Hence,

$$\left| P_{n-1}f(0) - p(0)^T \mathbf{V}^T \tilde{y} \right| = \left| (\mathbf{V} p(0))^T (y - \tilde{y}) \right|. \quad (2.49)$$

By Hölder's inequality,

$$\left| (\mathbf{V} p(0))^T (y - \tilde{y}) \right| \leq \|\mathbf{V} p(0)\|_1 \|y - \tilde{y}\|_\infty. \quad (2.50)$$

From Lemma 4, and from the assumptions on the distance between y and \tilde{y} ,

$$\|\mathbf{V} p(s)\|_1 \|y - \tilde{y}\|_\infty \leq \left(\frac{2}{\pi} \log(n+1) + 1 \right) \frac{\epsilon}{\frac{2}{\pi} \log(n+1) + 1} = \epsilon \quad (2.51)$$

with probability $\Pr = (1 - \delta/n)^n$. In fact, since the probability of each component \tilde{y} exceeding the specified distance is δ/n , by the union bound the total probability of at least one component exceeding this distance is less than $n \times (\delta/n) = \delta$. Thus, the inequality is satisfied with probability $\Pr \geq 1 - \delta$. This completes the proof. □

Theorem 5 is what suggests that our interpolation approach may have the potential to achieve accuracy improvements without increasing costs compared to standard Trotter. It tells us that the error in Trotter data can be as large as the error of the final estimate up to a factor which is logarithmically small in the number of interpolation points, and therefore these data \tilde{y}_i can be computed "cheaply enough." Thus, Theorem 5 is plays an important role in the proofs of Appendices B and E, and in the corresponding results presented below.

In the applications to Gaussian phase estimation we consider in Section 3.2.3, we will need a probabilistic statement of well-conditioning that takes into account confidence intervals for the data. This is supplied by the following lemma.

Lemma 6. *Let $\{Y_j\}$ be normally distributed random variables for the n Chebyshev interpolation points, with central values $\mu_j = f(s_j)$ and variances σ_j . Then the random variable $R_{n-1}(s)$ associated with the polynomial interpolant at value s is normally distributed with mean $P_{n-1}(s)$ and variance σ_R^2 bounded as*

$$\sigma_R^2(s) \leq 2 \max_j (\sigma_j^2). \quad (2.52)$$

Proof. The variable $R_{n-1}(s)$ is given by $p(s)^T \mathbf{V} Y$, with $Y = (Y_1, Y_2, \dots, Y_n)$. By linearity, the expectation value $E(R_{n-1}(s)) = p(s)^T \mathbf{V} E(Y) = p(s)^T \mathbf{V} y = P_{n-1}(s)$ as claimed, where $y = (f(s_1), \dots, f(s_n))$.

To bound the variance, observe that

$$\begin{aligned} \sigma_R^2(s) &= \text{Var}(R_{n-1}(s)) \equiv \text{Cov}(R_{n-1}(s), R_{n-1}(s)) \\ &= \sum_{i,j,k,l} \text{Cov}\left(p_i(s) (\mathbf{V}^{-1})_{ij} Y_j, p_l(s) (\mathbf{V}^{-1})_{lk} Y_k\right) \\ &= p^T(s) \mathbf{V}^{-1} \Sigma^2 \mathbf{V} p(s) \end{aligned} \quad (2.53)$$

$$= \|\Sigma \mathbf{V} p(s)\|_2^2, \quad (2.54)$$

where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$, and the last inner product was identified as a squared two norm. We can bound this using the spectral norm.

$$\sigma_R^2(s) \leq \|\Sigma\|^2 \|p(s)\|_2^2 \leq 2 \max_j \sigma_j^2 \quad (2.55)$$

□

At first, it may seem that the results of this lemma are quite restrictive, only applying to unbiased, normally distributed estimates of the function values $f(s_i)$. However, we will find occasion to apply this lemma in the context of the Gaussian phase estimation considered in Section 3.2.3.

3 Interpolation for Eigenvalues

3.1 Approach 1: Error Analysis through Perturbation Theory

Having laid the groundwork for well-conditioned polynomial interpolation, we apply our framework to the task of phase estimation. The idea is to perform logarithmically many phase estimation experiments evaluated at the Chebyshev nodes 2.17. We then bound the error using the interpolation theory of the previous section. The following theorem bounds the performance of such an algorithm.

Theorem 7. *Let $H : \mathbb{R} \mapsto \mathbb{C}^{2^n \times 2^n}$ be a Hamiltonian of the form $H(t) = \sum_{j=1}^m a_j(t) H_j$ for all $t \in \mathbb{R}$, where each H_j is a Hermitian matrix and $a_j(t)$ is real valued and in C^n for positive integer n . For each $t \in \mathbb{R}$, let $|\ell(t)\rangle$ be an eigenstate of $H(t)$ with eigenvalue*

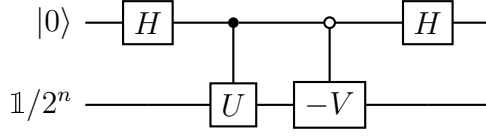


Figure 1: Quantum circuit for computing a normalized Frobenius distance $d(U, V)$ between unitaries U, V , where the probability of measuring zero on the first qubit is $d(U, V)^2$. By taking U as a single Trotter step and V as the s -dependent product formula of equation (2.7), polynomial extrapolation can be performed to estimate the true Trotter error.

$\lambda_\ell(t)$, and assume oracular preparation of $|\ell(t)\rangle$. Further, assume that there exists a minimum gap $\gamma(t) > 0$ such that for all $\ell' \neq \ell$ and $t > 0$, $|\lambda_\ell(t) - \lambda_{\ell'}(t)| \geq \gamma(t)$. Finally, that t is sufficiently small such that the assumption (2.11) of Lemma 1 holds. It is then possible to use an n -point polynomial interpolation formula over the $2k$ -th order Suzuki Trotter formula to estimate $\lambda_\ell(0)$, within error ϵ and failure probability at most $1/3$, using a number of operator exponentials and queries that is bounded above by

$$\tilde{O}\left(\frac{m^2(25/3)^k \max \|H_i\|(1 + \Gamma)}{\epsilon}\right),$$

where

$$\Gamma := \max_{t, p=1, \dots, n} \left(\frac{p}{n}\right) \left(\frac{ke^2(5/3)^k m \max_{l \in [1, m]} \|H_l\|}{\gamma(t)}\right)^{1/p}$$

and $n \in \tilde{O}\left(\log\left(mk(5/3)^k \max_i \|H_i\|(1 + \Gamma)\right) + \log 1/\epsilon\right)$.

The proof of this theorem is technical and relegated to Appendix B. As a brief sketch, the proof proceeds by using perturbation theory to evaluate the derivatives of the eigenvalues and eigenvectors of the effective Hamiltonian \tilde{H}_s , as well as multiple applications of the triangle inequality and similar linearizing approximations.

The above result applies generically to any piecewise analytic Hamiltonian $H(t)$. To give a better intuition for how it could be used, let us focus our attention on a phase estimation protocol. There are multiple phase estimation procedures that can be employed, but the shared basic idea is that one provides a quantum state $|\psi\rangle = \sum_j c_j(t) |j(t)\rangle$ and performs a series of evolutions of the form $e^{-iH(t)t}$ to this state to yield an estimate of one of the eigenvalues, $\exp(-i\lambda_l(t)t)$. Each eigenvalue is randomly sampled with probability $|c_j(t)|^2$. We will usually demand that the variance of this estimate is bounded above by t^2 and that the expected error is at most ϵ for such a procedure, but it is also common for the accuracy guarantees to be given in terms of a probability of failure.

3.1.1 Application: Estimation of Trotter Error

One important capability of our method is the estimation of error in ST formulas. This is important because existing error bounds are typically not tight, and leading order

expansions for the error are prohibitively expensive even for short evolutions [18]. We address this via a method for computing a distance metric d between two unitaries U, V . This metric is related to the Frobenius distance $\|U - V\|_F$, where

$$\|A\|_F := \sqrt{\text{Tr}A^\dagger A} \quad (3.1)$$

is the Frobenius norm. Specifically, we fix an ST order k and let \tilde{U}_s be as in equation (2.7). Our method applies to any product formula, but our theoretical analysis restricts us to consider ST formulas. To estimate the Trotter error, we compute

$$d(\tilde{U}_1(t), \tilde{U}_s(t)) := \frac{1}{\sqrt{2^n}} \left\| \frac{\tilde{U}_1(t) - \tilde{U}_s(t)}{2} \right\|_F \quad (3.2)$$

using the circuit of Figure 1. Observe that $\tilde{U}_1 = S_{2k}$ is a single Trotter step, while

$$\lim_{s \rightarrow 0} \tilde{U}_s(t) = U(t) \quad (3.3)$$

is the exact time evolution operator. Therefore, by estimating

$$d(S_{2k}(t), e^{-iHt}) = \lim_{s \rightarrow 0} d(\tilde{U}_s(t), \tilde{U}_1(t)), \quad (3.4)$$

we obtain a metric of the Trotter error. This metric $d \equiv d(\tilde{U}_1(t), \tilde{U}_s(t))$ corresponds to a root mean square of the singular values of $S_{2k}(t) - \tilde{U}_s(t)$. That is,

$$d(U, V) = \sqrt{\frac{1}{2^n} \sum_{j=1}^{2^n} (\sigma_j/2)^2} \quad (3.5)$$

for singular values $\sigma_j \in [0, 2]$ of $U - V$. It has the following relation with the spectral distance.

$$\frac{1}{\sqrt{2^n}} \left\| \frac{U - V}{2} \right\| \leq d(U, V) \leq \left\| \frac{U - V}{2} \right\| \quad (3.6)$$

Both the lower and upper bounds are tight generically. The upper bound is tight in the case that all singular values are equal. The normalized Frobenius norm is best interpreted as an average error over all "principal directions." Thus, though not a worst-case error, d still provides a useful characterization of the Trotter error.

Per usual, we consider an extrapolation of the normalized Frobenius distance d to $s = 0$.

Corollary 8. *Under the assumptions of Theorem 7, there exists a quantum algorithm that can compute the quantity $d(S_{2k}(t), e^{-iHt})$ for $H = \sum_{j=1}^m H_j$ with ϵ error and failure probability (also) ϵ using a number of operator exponentials N_{exp} of the H_j satisfying*

$$N_{\text{exp}} \in \tilde{O} \left(\frac{m^2 (25/3)^k \max \|H_i\|}{\epsilon} \right).$$

Proof. Consider the circuit of Figure 1, with $U = \tilde{U}_s$ and $V = \tilde{U}_1$. Let SELECT refer to the control operations in this LCU-type circuit. We first demonstrate that the probability of measuring the first qubit to be zero is $\|S_{2k}(t) - \tilde{U}_s(t)\|_F^2 / (2^{n+2})$. This fact follows from the analysis of the LCU lemma [9]; however, that analysis is typically performed for pure states, so we generalize it here. Let $|0\rangle\langle 0| \otimes \rho$ be the input state of the circuit. Then the gate operations proceed as follows.

$$\begin{aligned}
|0\rangle\langle 0| \otimes \rho &\mapsto_H \sum_{ij} \frac{|i\rangle\langle j|}{2} \otimes \rho \\
&\mapsto_{\text{SELECT}} \sum_{ij} \frac{|i\rangle\langle j|}{2} \otimes (\tilde{U}_s(t)^i (-\tilde{U}_1(t))^{1-i}) \rho (\tilde{U}_s(t)^j (-\tilde{U}_1(t))^{1-j})^\dagger \\
&\mapsto_H \sum_{ijkl} (-1)^{ik+j\ell} \frac{|k\rangle\langle \ell|}{4} \otimes (\tilde{U}_s(t)^i (-\tilde{U}_1(t))^{1-i}) \rho (\tilde{U}_s(t)^j (-\tilde{U}_1(t))^{1-j})^\dagger := \sigma
\end{aligned} \tag{3.7}$$

We then have that the probability of measuring the first qubit to be zero is

$$\text{Tr}(|0\rangle\langle 0| \otimes \rho \sigma) = \frac{1}{4} \text{Tr} \left((\tilde{U}_s(t) - \tilde{U}_1(t)) \rho (\tilde{U}_s(t) - \tilde{U}_1(t))^\dagger \right). \tag{3.8}$$

Taking $\rho = 1/2^n$ in particular yields

$$\begin{aligned}
\frac{1}{4} \text{Tr} \left((\tilde{U}_s(t) - \tilde{U}_1(t)) \rho (\tilde{U}_s(t) - \tilde{U}_1(t))^\dagger \right) &= \frac{\text{Tr} \left((\tilde{U}_s(t) - \tilde{U}_1(t)) (\tilde{U}_s(t) - \tilde{U}_1(t))^\dagger \right)}{2^{n+2}} \\
&= \frac{\|\tilde{U}_s(t) - \tilde{U}_1(t)\|_F^2}{2^{n+2}}
\end{aligned} \tag{3.9}$$

Thus, the probability of measuring 0 on the first qubit gives a normalized Frobenius distance between the two operators. The cost of doing this is $O(1)$ queries to the underlying ST formulas, each of which boils down to $O(m5^k)$ operator exponentials.

Using Amplitude Estimation on the target, we can construct an operator W such that the eigenvalues of W within the subspace supporting the initial state are of the form

$$\lambda(W) = \exp \left\{ \pm i \sin^{-1} \sqrt{\frac{\|\tilde{U}_s(t) - \tilde{U}_1(t)\|_F^2}{2^{n+2}}} \right\}. \tag{3.10}$$

We then can invoke Theorem 7 to show that the number of exponentials needed to learn the extrapolated phase, within error ϵ' with probability greater than $2/3$, satisfies

$$N_{\text{exp}} \in \tilde{O} \left(\frac{m^2 (25/3)^k \max \|H_i\| \log^{5/2}(\max_i \|H_i\| / \epsilon')}{\epsilon'} \right). \tag{3.11}$$

The remaining question is how small ϵ' must be to guarantee that the error in the normalized Frobenius distance is at most ϵ . Let $\hat{\phi}$ denote the estimate of the phase

returned by our protocol, which has error at most ϵ' . Our estimate \hat{d} of the distance metric is then related to the phase in equation (3.10) as

$$\hat{d} = \sin(\hat{\phi}). \quad (3.12)$$

Thus we have that

$$|\hat{d} - d| = |\sin \hat{\theta} - \sin \theta| \leq |\hat{\theta} - \theta| \leq \epsilon'. \quad (3.13)$$

Choosing $\epsilon' = \epsilon$, then, is sufficient to ensure the desired tolerance as given in the lemma. \square

As mentioned, d is not a worst-case simulation error, but it may be a good approximation in typical situations. To get a true upper bound, we could employ the relation $\|A\| \leq \|A\|_F$, but estimating $\|U - V\|_F$ directly from d requires precision exponential in n . Moreover, the upper bound would not be very tight unless d were already quite close to $\|U - V\|$.

In summary, our estimates should be better when the spread of the eigenvalues in the error operator is not too large. We formalize this in the following theorem.

Corollary 9. *Let $\delta U^2 = |\tilde{U}_s(t) - \tilde{U}_1(t)|^2$, and suppose that*

$$\sigma^2(\delta U^2) = \text{Tr}(\delta U^4)/2^n - \text{Tr}(\delta U^2/2^n)^2 \leq \xi$$

for some $\xi > 0$. Then each eigenvalue λ^2 of δU^2 satisfies

$$\Pr \left(\left| \lambda^2 - \frac{\|\tilde{U}_s(t) - \tilde{U}_1(t)\|_F^2}{2^n} \right| \geq k\sqrt{\xi} \right) \leq \frac{1}{k^2}. \quad (3.14)$$

for all $k > 0$. Further, even with no such promise about the variance, the following weaker bound holds.

$$\Pr \left(\left| \lambda^2 - \frac{\|\tilde{U}_s(t) - \tilde{U}_1(t)\|_F^2}{2^n} \right| \geq k \frac{\|\tilde{U}_s(t) - \tilde{U}_1(t)\|_F^2}{2^n} \right) \leq 1/k \quad (3.15)$$

Proof. The result follows from the Chebyshev and Markov inequalities. \square

Corollary 8 and Corollary 9 show that we can use our interpolation procedure to the estimate largest eigenvalue of the error operator. In particular, let the probability of an eigenvalue being greater than the estimate be $O(1/2^n)$. Then the Chebyshev bound (3.14) implies that it suffices to take $k \in O(\sqrt{2^n})$. Thus, with high probability, all of the eigenvalues for the square of the error operator will be at most $4\hat{d}^2 + O(\sqrt{\xi 2^n})$. Thus if $\xi \in o(\hat{d}^4/2^n)$ then the estimate yielded by this procedure will also estimate the spectral norm.

3.2 Approach 2: Error Analysis through Bernstein Ellipses and Analyticity

In this section, we propose an alternative analysis for eigenvalue estimation, as well as a specific approach to phase estimation. We first prepare each effective eigenstate through the procedure in [36], except using one single qubit the semi-classical QFT, summarized in Theorem 15. Then we perform a new Gaussian phase estimation (Theorem 16) on said eigenstate and then interpolate the estimates. The main result is summarized in Theorem 17.

To analyze the error and complexity of this approach, we use a different formalism that relies on complex analyticity [37], which allows us to estimate the convergence rate of the interpolation in terms of range of analyticity rather than and not derivatives of the effective Hamiltonian. This is beneficial since the bounds using derivatives might become unmanageable near level crossings (See Appendix B).

3.2.1 Ancillary Lemmas

To begin, we need to introduce some notions from complex analysis. For each $\rho > 1$, let $B_\rho \subset \mathbb{C}$ be the Bernstein ellipse, which is an ellipse with foci at ± 1 and semimajor axis $(\rho + \rho^{-1})/2$. The following lemma bounds the Chebyshev interpolation error for analytic functions on B_ρ .

Lemma 10. *Let $f(z) \in \mathbb{C}$ be an analytic function on B_ρ , and suppose $C \in \mathbb{R}_+$ is an upper bound such that $|f(z)| \leq C$, for all $z \in B_\rho$. Then the Chebyshev interpolation error on $[-1, 1]$ satisfies*

$$\|f - P_n f\|_\infty \leq \frac{4C\rho^{-n}}{\rho - 1}$$

for each degree $n > 0$ of the interpolant through the $n + 1$ Chebyshev nodes.

Proof. Theorem 8.2 of Ref. [37]. □

Lemma 10 shows that the interpolation error shrinks exponentially in n . We would like to apply this lemma to analyze the eigenvalues $\lambda(z)$ of the effective Hamiltonian \tilde{H}_z , which is a continuation of \tilde{H}_s to the complex plane. However, we need to understand the domain under which $\lambda(z)$ is analytic, i.e. free of level-crossings.

To characterize this domain, we first utilize a result from Bauer and Fike [38] which bounds the shift in the eigenvalues under a shift in the operator.

Lemma 11. *Let A be a normal matrix with eigenvalues $\{\lambda_i\}$. Then, if λ is an eigenvalue of a matrix B , there exists an eigenvalue λ_k of A such that*

$$|\lambda - \lambda_k| \leq \|B - A\| \tag{3.16}$$

for at least one eigenvalue λ_k of A , where $\|\cdot\|$ is the spectral norm.

In order to determine the radius of analyticity, we assume a specific form for the upper bound on $\|H - \tilde{H}_z\|$, where \tilde{H}_z is the analytic continuation of \tilde{H}_s to the complex plane. Here is the resulting theorem

Lemma 12. Let $\lambda(\tau)$ be an eigenvalue of \tilde{H}_z , where $\tau = zt$ for $z \in \mathbb{C}$ and $t \in \mathbb{R}^+$, and γ_0 be a lower bound on the spectral gap of $\tilde{H}_0 = H$. Let $\alpha, \beta \in \mathbb{R}^+$ be constants such that

$$\|H - \tilde{H}_z\| \leq \frac{\alpha|\tau|^p}{(p+1)!} e^{\beta|\tau|}. \quad (3.17)$$

Then $\lambda(\tau)$ is analytic on any origin-centered disc of radius r provided that $r \leq r_{\max}$, with

$$r_{\max} := \frac{p}{\beta} W_0 \left(\frac{\beta}{p} \left(\frac{\gamma_0(p+1)!}{2\alpha} \right)^{1/p} \right). \quad (3.18)$$

Here W_0 is principal branch of the Lambert W function.

Proof. See Appendix D.1 □

The above lemma gives an upper bound r_{\max} on the radius of analyticity. From this, assuming $r_{\max} > 1$, we can calculate the largest possible B_ρ by equating the semimajor axis with the maximum radius

$$(\rho_{\max} + \rho_{\max}^{-1})/2 = r_{\max} \quad (3.19)$$

which has solution

$$\rho_{\max} = r_{\max} + \sqrt{r_{\max}^2 - 1}. \quad (3.20)$$

The following lemmas give a clue as to what α and β should look like in terms of commutators between the terms of H and also with respect to its overall spectral norm. However, first, we introduce a bound on the error on the effective total Hamiltonian evolution in terms of the "instantaneous" Hamiltonian error, \mathcal{E} , which is defined through the complex-time Schrödinger equation. Let $S_p(\tau)$ be a p th order product formula approximating $\exp(iH\tau)$, such that

$$\frac{d}{d\tau} S_p(\tau) = i(H + \mathcal{E}(\tau))S_p(\tau). \quad (3.21)$$

For this purpose, we also introduce the "accumulated" error

$$E(\tau) = \log S_p(\tau) - i\tau H, \quad (3.22)$$

where $\log S_p(\tau)$ can be defined through the complex-time Magnus expansion (See [39] for compilation of proofs, complex-time extension is straight-forward)

$$\log S_p(\tau) = \Omega = i \int_{\mathcal{P}} \sum_{n=0}^{\infty} \frac{B_n}{n!} \text{ad}_{\Omega}^n(H + \mathcal{E}(\tau_1)) d\tau_1 \quad (3.23)$$

where \mathcal{P} is a path going from $\tau_1 = 0$ to $\tau_1 = \tau$. This, like the real-time Magnus expansion, converges provided $\|\Omega\| \leq \pi$. With this, we obtain an upper bound on $\|E(\tau)\|$ stated through the following lemma:

Lemma 13. *Given the error definition*

$$E(\tau) = \log S_p(\tau) - i\tau H$$

for complex τ , we can bound its norm through

$$\|E(\tau)\| \leq \frac{5}{2} \int_0^{|\tau|} \|\mathcal{E}(s\tau/|\tau|)\| ds,$$

provided that $\max_{|\tilde{\tau}| \leq |\tau|} \|E(\tilde{\tau})\| \leq 15/8$, $|\tau| \leq 1/8$, and $\int_0^{|\tau|} \|\mathcal{E}(s\tau/|\tau|)\| ds \leq 1/8$.

Proof. See Appendix D.2 □

Now, we bound the norm for $\mathcal{E}(\tau)$ in Appendix D.3 and make use of $\|H - \tilde{H}_z\| \leq \|E(\tau)\|/|\tau|$.

Lemma 14. *Let \tilde{H}_z be the effective Hamiltonian associated with a complex-time p th order product formula $S_p(\tau)$, where $\tau = zt$ with $z \in \mathbb{C}$ and $t \in \mathbb{R}_+$, which is analytically continued to an open neighborhood containing $[-1, 1]$. The norm of operator error can upper bounded as*

$$\begin{aligned} \|H - \tilde{H}_z\| \leq \frac{5}{2} \sum_{(v,m)} \alpha_{\text{comm}} \left(a_{(\Upsilon,m)} H_{\pi_{\Upsilon}(m)}, \dots, a_{(v,m+1)} H_{\pi_v(m+1)}, a_{(v,m)} H_{\pi_v(m)} \right) \\ \cdot \frac{|\tau|^p}{(p+1)!} e^{2|\tau| \sum_j^m \|H_j\|} \end{aligned}$$

provided that $|\tau| \|H - \tilde{H}_z\| \leq 15/8$, $|\tau| \leq 1/8$, and

$$\sum_{(v,m)} \alpha_{\text{comm}} \left(a_{(\Upsilon,m)} H_{\pi_{\Upsilon}(m)}, \dots, a_{(v,m+1)} H_{\pi_v(m+1)}, a_{(v,m)} H_{\pi_v(m)} \right) \frac{|\tau|^p}{(p+1)!} e^{2|\tau| \sum_j^m \|H_j\|} \leq \frac{1}{20}.$$

Here, $\alpha_{\text{comm}}(A_s, \dots, A_1, B) := \sum_{q_1 + \dots + q_s = p} \binom{p}{q_1 \dots q_s} \left\| \text{ad}_{A_s}^{q_s} \cdots \text{ad}_{A_1}^{q_1}(B) \right\|$.

While the above theorem is a useful theoretical bound, a problem arises because bounding the commutators is often impractical computationally. In such cases, α_{comm} can be upper bounded through the triangle inequality and submultiplicativity of the norm.

$$\alpha_{\text{comm}}(A_s, \dots, A_1, B) \leq \|B\| 2^p \left(\sum_{j=1}^s \|A_j\| \right)^p. \quad (3.24)$$

With these set of lemmas at our disposal, we can now provide a cost estimate for state preparation for all the ground states of \tilde{H}_{s_k} .

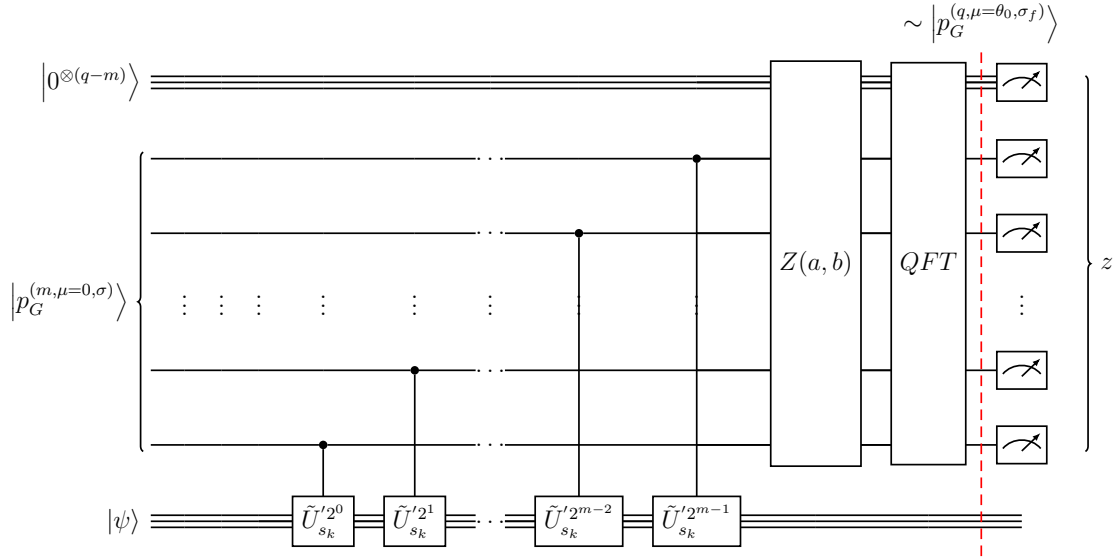


Figure 2: Circuit to implement a Gaussian m -qubit phase estimation algorithm with $(m-q)$ -qubits for spectral interpolation. The \tilde{U}'_{s_k} operator is evolved by an effective Hamiltonian \tilde{H}_{s_k} , with a Trotter step size ts_k (See Theorem 15) and corresponding evolution time $T'_k = e'_k \times dt = \Theta(ts_1)$.

3.2.2 State Preparation Cost

Theorem 15 (State preparation cost). *Let $H = \sum_{j=1}^m H_j$ satisfy $\sum_{j=1}^m \|H_j\| \leq 1$, and let $\gamma \in \mathbb{R}_+$ be a lower bound on the spectral gap of H . Let s_1, \dots, s_n be the n th degree Chebyshev nodes on $[-1, 1]$ for even n . Then there exists an algorithm to prepare the ground state of \tilde{H}_{s_k} to precision ϵ using controlled queries of $\tilde{U}'_{s_k} = (S_p(ts_k))^{e'_k}$, where $t \in \Theta(\gamma^{1/p})$ and $e'_k = \text{sgn } s_k \lceil s_1/|s_k| \rceil$, with a total gate count*

$$C_{\text{prep}} \in O\left(\frac{\log(1/\epsilon) \log \log(1/\epsilon)}{\gamma^{1+1/p}}\right)$$

and a single auxiliary qubit.

Proof. See Appendix D.5. □

Now that the cost of state preparation is covered, we assume an exact state preparation for our next gate count upper bound for eigenvalue estimation. However, first, we will introduce a new phase estimation algorithm as one of our methods for eigenvalue estimation.

3.2.3 Interlude: Gaussian Quantum Phase Estimation

The first step in the phase estimation protocol involves preparing a Gaussian distribution in the ancillary register of m qubits. Let $p(w; \sigma)$ be the probability density function

for a Gaussian distribution with zero mean ($\mu = 0$) and variance σ^2 .

$$p(w; \sigma) := \frac{1}{\sigma\sqrt{2\pi}} e^{-w^2/(2\sigma^2)} \quad (3.25)$$

Define

$$|p_G(\sigma, T, m)\rangle := \sum_{x=-(2^{m-1}-1)}^{2^{m-1}-1} \sqrt{\frac{p(xT; \sigma)}{\mathcal{N}(\sigma, T, m)}} |x\rangle \quad (3.26)$$

where $T > 0$ plays the role of a sampling rate, and $\mathcal{N}(\sigma, T, m) > 0$ is the normalization constant for the discrete, truncated Gaussian.

$$\mathcal{N}(\sigma, T, m) := \sum_{x=-(2^{m-1}-1)}^{2^{m-1}-1} p(nT; \sigma). \quad (3.27)$$

An exact preparation of $|p_G\rangle$ can be achieved through the methods proposed in Ref. [40]. However, within a target error, we can prepare a coarser Gaussian distribution and then perform an upsampling through QFT, zero-padding, and then QFT⁻¹. This is an established method for upsampling/interpolation of discrete signals in classical discrete signal processing (DSP), and was introduced for quantum distribution preparation in Ref [41]. In Figure 2 we illustrate the circuit used for this method. The operator Z plays the role of zero-padding ubiquitous to classical DSP for interpolation in the conjugate space. The errors introduced by truncation and finite sample rate in the time domain are estimated in Appendix C, and can be summarized with the following theorem.

Theorem 16. *Let q, m be positive integers such that $q \geq m$. Let $T, \sigma > 0$ be chosen such that $\sigma/T = \Theta(\sqrt{2^m})$, and define the Fourier conjugates $F := 1/(2^q T)$ and $\sigma_f := 1/(4\pi\sigma)$. Finally, let*

$$X(f) := \mathcal{F}\{\sqrt{p(\cdot; \sigma)}\}(f) := \int_{\mathbb{R}} \sqrt{p(t; \sigma)} e^{-2\pi i t f} dt = \sqrt{\sqrt{2\pi}\sigma_f} p(f; \sigma_f) \quad (3.28)$$

denote the Fourier transform of the (root of the) Gaussian p . Then the spectral norm error of the prepared wave function in Fourier domain is

$$\left\| \sum_k^{2^q} \frac{X(kF)}{\sqrt{\mathcal{N}(\sigma_f, F, q)}} |k\rangle - \text{QFT} |p_G(\sigma, T, m)\rangle \right\| \in O\left(\frac{2^{m/2}}{e^{\Omega(2^m)}}\right). \quad (3.29)$$

Using these errors in the approximate state constructed in Fourier domain, we can perform a sequence of unitaries controlled on the Gaussian window. This approach is similar in spirit to the Kaiser-window approach taken in [42]. However, here we use a Gaussian window, which by inverting the conjugate spaces labels in Theorem 16, we know we can prepare efficiently with a cost $O(\text{polylog}(1/\varepsilon)/\sigma)$, where ε is the vector error on the window state.

3.2.4 Phase Estimation Cost

Via the above theorem, we can now bound the cost for extrapolating eigenvalues obtained using Gaussian phase estimation. Through Theorem 5 and Lemma 6, we estimate the effects of uncertainty propagation that interpolation has on the final interpolant for Gaussian and one-qubit phase estimation. The result goes as follows:

Theorem 17 (Phase estimation cost). *The gate count of estimating the eigenvalues of \tilde{H}_{s_k} using Gaussian phase estimation is*

$$C_{\text{est,gpe}} \in O\left(\frac{\log(1/\epsilon)}{\sigma_P}\right),$$

where σ_P is the standard deviation on the interpolated observable at $s = 0$. The number of auxiliary qubits required is $O(\log(1/\sigma_P))$. Alternatively, using a single-ancillary-qubit approach,

$$C_{\text{est,1-qubit}} \in O\left(\frac{\log(1/\epsilon) \log \log(1/\epsilon)}{\epsilon}\right),$$

where $\epsilon > 0$ is such that $\max_s |\lambda(s) - P_{n-1}\tilde{\lambda}(s)| \leq \epsilon$ and $P_{n-1}\tilde{\lambda}$ is the (unique) $(n-1)$ -degree polynomial interpolating the data $\tilde{\lambda}_i$ at the n interpolation points s_i .

The bound for $C_{\text{est,gpe}}$ can also be cast in terms of a confidence interval, $\epsilon = w\sigma_P$, around the mean, which introduces an error rate that decreases super-exponentially with w .

4 Interpolation for Expectation Values

We now consider the application of Chebyshev interpolation to estimate expectation values, a fundamental task in quantum computation. The setting is as follows: given a quantum state ρ and observable O , the expectation value is given by $\langle O \rangle = \text{Tr}(\rho O)$. We evolve our system according to a $2k$ -th order ST formula \tilde{U}_s given by (2.7). The time evolved expectation values of interest is captured by the function

$$f(s) := \frac{\text{Tr}(\rho O_s(t))}{\|O\|} \tag{4.1}$$

where $O_s(t)$ is given by equation (2.13). We've normalized the expectation values by $\|O\|$ because the relative error is a useful and natural metric, and also the normalized operators may be block encoded for amplitude estimation. Alternatively, we simply restrict our attention to normalized observables with $\|O\| = 1$. The interpolation algorithm we propose can be summarized as follows.

1. Given Hamiltonian H , simulation time t , and tolerance ϵ for the estimate of $\langle O(t) \rangle / \|O\|$, choose the appropriate interpolation interval $[-a, a]$ and an even number n of Chebyshev nodes. We neglect the cost of this step. The error analysis we will perform subsequently will inform the choices of a and n .
2. Compute estimates \tilde{y}_i of the expectation values $\langle O_{s_i}(t) \rangle$ for each s_i with $i = 1, \dots, n/2$, to an accuracy depending on ϵ and n . We will assume this step is done with Iterative Quantum Amplitude Estimation (IQAE) [43], a recent approach to amplitude estimation that exhibits low quantum overhead. Our metric of cost is the number of H_j exponentials executed on a quantum circuit, where $H = \sum_j H_j$. Note that by symmetry, we need not compute \tilde{y}_i for $i > n/2$. We have $f(s_i) = f(s_{n-i+1})$ for all $i \in \{1, \dots, n\}$.
3. Perform the polynomial fit $\tilde{P}_{n-1}f$ through the points (s_i, \tilde{y}_i) using a Chebyshev expansion (2.26). Note that $\tilde{P}_{n-1}f$ will automatically be even. This fit is well-conditioned, and we neglect the cost of this step.
4. Evaluate the $\tilde{P}_{n-1}f(0)$ to be our estimate of $\langle O(t) \rangle$.

To summarize, one performs amplitude estimation to acquire the time evolved expectation value at each Chebyshev node, then performs a polynomial interpolation of the data. The estimate is the value at $s = 0$.

Given an even set of Chebyshev nodes $\{s_1, \dots, s_n\}$, and making use of Lemma 2, the interpolation error E_{n-1} assuming perfect data points is given by

$$|E_{n-1}(0)| \leq \frac{|\text{Tr}(\rho \partial_s^n O_s(t))|}{\|O\| n!} \prod_{i=1}^n |s_i| \leq \max_{s \in [-a, a]} \frac{\|\partial_s^n O_s(t)\|}{\|O\|} \left(\frac{a}{2n}\right)^n. \quad (4.2)$$

With a suitable bound on $\partial_s^n O(t)$, we can provide an upper bound on the interpolation error at $s = 0$. This bound is provided by the following lemma. In what follows, it will be helpful to define the parameter

$$c := k(5/3)^k m \max_{l \in [1, m]} \|H_l\| t \quad (4.3)$$

for ease of notation. This parameter is proportional to the the Hamiltonian size and the "total Trotter time," meaning the sum of all the forward and backward time steps, in absolute value, for a $2k$ th ST formula.

Lemma 18 (Error extrapolation for time-evolved observables). *Under the conditions of Lemma 1 ($ca \leq \pi/20$), the following bounds holds on the Trotterized evolution $O_s(t)$ with step parameter $s \in (0, a]$:*

1. for $c > n$ we have that

$$\frac{\|\partial_s^n O_s(t)\|}{\|O\|} < \left(c \sqrt{e^3(1 + \sqrt{8/\pi}e^2)} \right)^{2n}$$

which gives an interpolation error

$$|E_{n-1}(0)| < \left(129 \frac{c^2 a}{n}\right)^n.$$

2. For $c \leq n$, we have

$$\frac{\|\partial_s^n O_s(t)\|}{\|O\|} \leq \sqrt{\frac{2n}{\pi}} \left(\frac{e^4 c}{2}\right)^n n! e^{4cc^2 \sqrt{2/\pi}}$$

giving an interpolation error

$$|E_{n-1}(0)| \leq 2\sqrt{2n} (6ca)^n e^{24c}.$$

The proof of this lemma is a tedious exercise in repeated in the combinatorics of large derivatives and the triangle inequality, and is left to Appendix E. Note that once the derivative bound holds, the interpolation error bound follows immediately from Lemma 2.

One motivation for these bounds is deriving asymptotic expressions for the algorithmic complexity. The following theorem gives an asymptotic query complexity for the number N_{exp} of Trotter exponentials $\exp(-iH_j\tau)$.

Theorem 19. *Let $O(t) = U^\dagger(t)OU(t)$ be a time-evolved observable under a Hamiltonian $H = \sum_{l=1}^m H_l$ on n qubits, so that $U(t) = e^{-iHt}$. Suppose there exists a $\gamma \in \mathbb{R}_+$ such that O/γ can be block encoded via a unitary U_{enc} by a state $|G\rangle$ on a set of L auxiliary qubits. Let ρ be a quantum state on n qubits, and suppose $\gamma/\|O\| \in O(1)$. Then, the number of exponentials N_{exp} required to estimate $\text{Tr}(\rho O(t))/\|O\|$ to precision ϵ with confidence $1 - \delta$ using a $2k$ order Suzuki Trotter formula satisfies*

$$N_{\text{exp}} \in \tilde{O}\left(c \max\{c, \log(1/\epsilon)\} \epsilon^{-1} \log(1/\delta)\right).$$

Here, \tilde{O} is big- O with multiplicative terms suppressed which are logarithmically smaller in $1/\epsilon$ and c . Moreover, the number of auxiliary qubits needed is $O(L)$.

We give a sketch of the proof. Given a choice of interval $[-a, a]$ and (even) number of interpolation points n , we have from Lemma 3 that the number of exponentials to perform evolutions for all Chebyshev nodes goes as

$$O\left(\frac{n \log n}{a}\right). \tag{4.4}$$

However, this is not the total cost since these circuits need to be repeated to perform the appropriate measurement protocols. Since O can be block encoded, the expectation value can be obtained via an amplitude estimation protocol. By the well-conditioning of Lemma 5, each data point needs to be within ϵ of the exact Trotter value, up to a logarithmic factor in n . This robustness is why our result maintains a $\tilde{O}(\epsilon^{-1})$ scaling.

In our proof, we assume the IQAE protocol is used, requiring only a single qubit overhead. The fractional queries for the noninteger timestep also require $O(1)$ overhead, meaning the total overhead is $O(L)$ due to the block encoding. To relate n and a to the required precision ϵ , simulation time t and Hamiltonian H , Lemma 18 can be used. Thus, we can relate N_{exp} to these basic parameters. We carry out the formal proof in Appendix E.

As advertised, we see there is a "near-Heisenberg" scaling of $1/\epsilon$, up to logarithmic factors. However, there is an unsavory quadratic scaling in the simulation time in cases without high accuracy demands. We believe this can be improved, because our approach us forces us to have n scale linearly in T . We believe this is overly pessimistic, and a $\log(\|H\| T)$ scaling is more likely. Finally, our results suggest the best performance for using low order formulas, since our bounds are strictly worse for increasing ST order k .

5 Numerical Experiments

In the previous sections, we presented theorems, with proofs in the appendices, that interpolation of Trotter data can lead to higher accuracies for eigenvalues and expectation values than Trotter alone. This section provides numerical evidence backing up our analytic claims, not only showing that improved scaling is possible, but also that high-order Trotter formulas need not always provide better error scaling when used in conjunction with interpolation. Specifically, we demonstrate this improved scaling for phase estimation using interpolation for second-order and fourth-order ST formulas. For our demonstration we consider the transverse Ising model of two spins.

$$H = -J(Z \otimes Z + g(X \otimes I + I \otimes X)) \quad (5.1)$$

This is a minimal example with just two non-commuting terms, yielding a nonzero Trotter error. Our aim will be to estimate the ground state energy of the effective Hamiltonian \tilde{H}_s in the limit $s \rightarrow 0$, and seek numerical evidence of the improved performance relative to low-order Trotter formulas that Theorems 15 and 17 suggest.

In order to avoid aliasing with our Fourier spectral methods, we must satisfy the bound

$$\left\| \tilde{H}_{s_k} \right\| \frac{t}{2\pi} |e'_k s_k| + \Delta_{\text{pad}} \leq 1, \quad (5.2)$$

where

$$e'_k = \text{sgn } s_k \left[\frac{s_1}{|s_k|} \right]. \quad (5.3)$$

The padding Δ_{pad} is there to suppress the probability leakage wrapping around the boundary. For simplicity, define a dimensionless effective Hamiltonian

$$\tilde{h}_s := \tilde{H}_s \frac{te'_s s}{2\pi}, \quad (5.4)$$

with $e'_s = \text{sgn } s \lceil s_1/|s| \rceil$, such that

$$\left\| \tilde{h}_{s_k} \right\| + \Delta_{\text{pad}} \leq 1. \quad (5.5)$$

This way, the spectrum of \tilde{h}_{s_k} lies within the domain $[-1/2 + \Delta_{\text{pad}}/2, 1/2 - \Delta_{\text{pad}}/2]$ and now is in line with the conventions of Fourier spectral methods in [36]. With this, we now define the normalized Hamiltonian

$$h = \tilde{h}_0. \quad (5.6)$$

Here t is sufficiently small as to fulfill Equation (5.5). We must also make sure that s is chosen according to the constraints of Section 3.2 such that there are no level crossings throughout the interpolation interval.

We now describe the results. Figure 3 displays the results of Chebyshev interpolation with second order Trotter formulas using various (even) numbers of points. For each \tilde{H}'_{s_k} , we numerically simulate the ground state preparation protocol to a state error, ϵ_{state} . This is efficient because the algorithm scales $O(\log 1/\epsilon_{\text{state}})$. We then simulate the Gaussian phase estimation on these states. Moreover, we have used zero-padding as described in the Section 3.2.3 in order to upsample the spectrum and be able to ignore digitization errors. The spectral upsampling through zero-padding is efficient since the cost scales like $O(q^2)$, where q is the number of total qubits after padding. The blue bands in these plots show an uncertainty equivalent to $0.01 \times \sigma_P(s)$. We have down-scaled the statistical uncertainties by a factor of a 100 such that the interpolation error changes are noticeable when we change the number of nodes.

Next, in Figure 4, we show the exact systematic error from our simulations as well as the upper bounds of the Bernstein ellipse analysis of Lemmas 14, 12, and 10. We see from this data that interpolation indeed improves the quality of the estimate of the energy as anticipated, even when relatively low-precision estimates are used at $s = 1$. As discussed previously, only positive values of s_k are computed due to reflection symmetry.

Finally, in Figure 5 we performed ground state energy estimation using second and fourth order formulas without interpolation, in order to compare the costs when plotted against the error. From the plots, we observe that beyond $n = 2$ the extrapolation methods already outperform using second-order product formulas alone. Specifically, we see clear indications from this data that the bias in the interpolated error (for large $1/\epsilon$) scales logarithmically for these Gaussian phase estimation experiments. This agrees with our expectations from Theorem 17 wherein the systematic errors from interpolating phase estimation experiments are expected to scale as $\text{polylog}(1/\epsilon)$. In contrast, the bias from a fixed experiment is expected to be $O(1/\epsilon^{3/2})$ and $O(1/\epsilon^{5/4})$ respectively.

It is worth noting that this does not violate the Heisenberg limit as here we are only showing the interpolation error, with data computed to machine precision. The exponential reduction in the uncertainty with the number of interpolation points considered

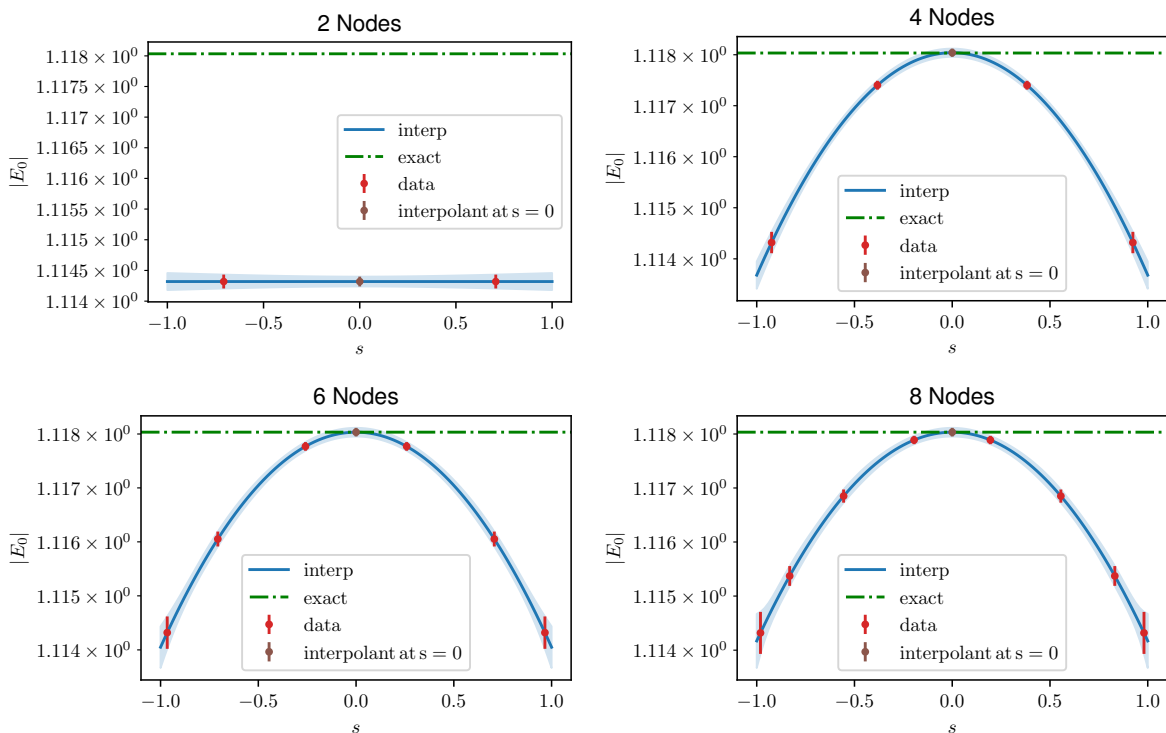


Figure 3: Estimation of the ground state energy of the two-spin transverse Ising model using interpolation with Gaussian phase estimation. The four plots, left to right and top to bottom, have $n = 2, 4, 6, 8$ interpolation points. In all, the second order Trotter formula was used for simulation. The statistical uncertainties on the data and the interpolants shown here are equivalent to 0.01 times their standard deviation. They are scaled this way so that interpolation error changes are noticeable.

is independently shown in Figure 4. Note that for this data we do not see an improvement from transitioning to higher-order formulas, which is consistent with our prior expectations. These results show that extrapolation tends to outperform transitioning to a higher-order product formula does not necessarily lead to an improved scaling for the problem of phase estimation.

6 Conclusion

In this work, we present a framework for achieving improved accuracy of Trotter simulations using polynomial interpolation. We apply our framework to the problems of eigenvalue and expectation value estimation, and derive query complexities assuming the Trotter data is computed quantumly. We find that our methods achieve a Heisenberg scaling of $\tilde{O}(1/\epsilon)$ up to polylogarithmic factors. This is an improvement relative to a single Trotter calculation, which would yield $\tilde{O}(1/\epsilon^{1+1/p})$ for order p formulas. For the low-order formulas used in practice, where p is not so large, this improvement may

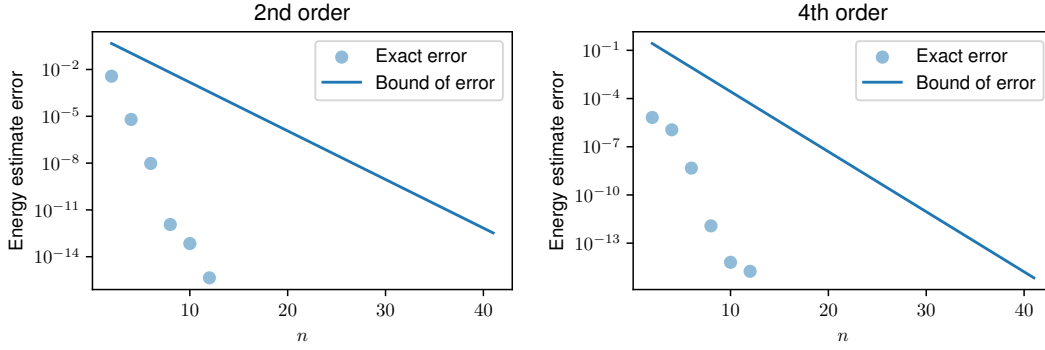


Figure 4: The systematic error coming from truncating the polynomial expansion. In blue we have the exact result for the error and in orange are the upper bounds for the error. On the left, we used a second-order formula and on the right a fourth-order one. As we can see from this log-log plot, the convergence is exponential as expected.

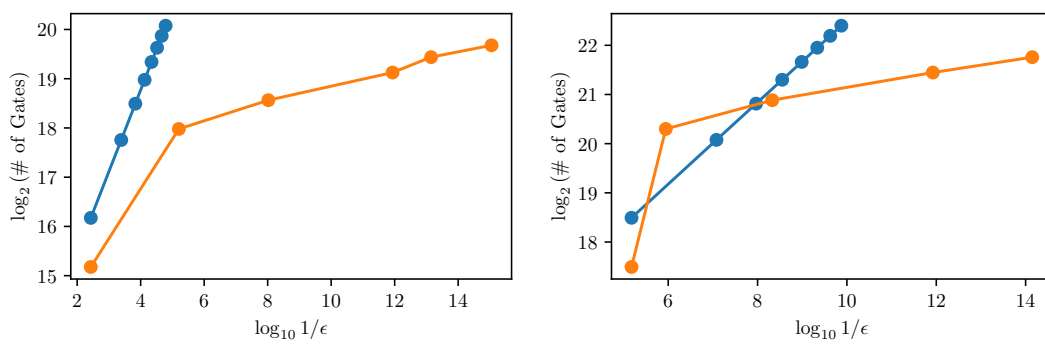


Figure 5: The gate cost from using second- (left) and fourth-order (right) formulas for ground state estimation plotted against the systematic error on the eigenvalue estimation. In blue, we have the cost of using single long-depth circuits using product formulas alone. In orange, we have the gate cost of using multiple shallower circuits and then using the extrapolation methods presented in this work.

be significant.

An interesting feature of polynomial interpolation in this context is that it provides a uniform approximation of the function $f(s)$ on the interpolation interval, thereby giving a characterization of the Trotter error for various step sizes. In contrast, similar techniques such as Richardson extrapolation only attempt to estimate $f(0)$. Whether this global characterization of the Trotter error can be usefully employed is a question we leave open.

There are several areas in which improvements in our analysis can likely be made. First, in most of our claims, we resort to fractional queries in order to produce step sizes at the values of the Chebyshev nodes. We suspect this machinery is not needed in practice, and that improved analysis, coupled with mathematical results on the robustness of the Lesbesgue constant to node perturbations, will show that interpolating at nearby points will be sufficient. Second, the quadratic dependence in simulation time that arises for expectation value interpolation in certain regimes is unsavory. The root cause of this in the analysis appears to be a linear relation between the simulation time and number of interpolation nodes, which seems overly cautious. Numerical studies beyond our own, as well as implementations on quantum hardware, may reveal a rosier picture than our naive bounds suggest. Third, our work does not consider the effects of imperfect operations. We can reasonably expect quantum operational errors to be more severe for the foreseeable future. In this case, the most salient effect is additional error on the data values that, following any mitigation procedures, cannot be arbitrarily diminished. We should expect that this will limit the interpolation accuracy to a certain threshold given roughly by the product $\epsilon_{\text{phys}} \times \Lambda$, where ϵ_{phys} is the typical error from imperfect operations and Λ is the Lesbesgue constant. Finally, our results do not show commutator scaling, though the method itself almost certainly does. For Hamiltonians with almost commuting terms, this would manifest in the data being nearly flat, resulting in an approximately straight line fit to the correct answer. Our difficulty in showing this expected relationship arises from the messy expressions for high order derivatives of the effective Hamiltonian, but this should not be a fundamental issue.

Our work shows that polynomial interpolation can be a powerful supplement to quantum subroutines using low-order Trotter simulations. Given that our results ameliorate the worst aspects of low order formulas, we believe that investigation of highly efficient low-order formulas for simulation is likely to be an increasingly important topic. Improvements that arise from the superior constant factors that low order formulas can provide may lead to a new generation of more efficient algorithms for simulating chemistry and physics using quantum computers.

7 Acknowledgements

We thank Christoph Lehner, Taku Izubuchi and Yuta Kikuchi for the insightful discussions. We also thank the reviewers for improving this article with specific feedback.

During the inception of this work G. R. was supported through Brookhaven National Laboratory (BNL), the Laboratory Directed Research and Development (LDRD) program No. 21-043, by Program Development Fund No. NPP PD 19-025, and by the US DOE Office of Nuclear Physics, under the Contract No. DE-SC0012704 (BNL). During later additions to the manuscript, G.R. and N.W. were supported by the US DOE National Quantum Information Science Research Centers, Co-design Center for Quantum Advantage (C2QA) under contract number DE-SC0012704. G.R. was also supported and employed by Zapata Computing, Inc. through part of the writing of the manuscript. J.W. is supported by the NSF Graduate Research Fellowship and U.S. Department of Energy grant No. DE-SC0023658.

References

- [1] S. Lloyd, *Universal quantum simulators*, *Science* **273** (1996) 1073.
- [2] M. Reiher, N. Wiebe, K.M. Svore, D. Wecker and M. Troyer, *Elucidating reaction mechanisms on quantum computers*, *Proceedings of the National Academy of Sciences* **114** (2017) 7555.
- [3] J.D. Whitfield, J. Biamonte and A. Aspuru-Guzik, *Simulation of electronic structure hamiltonians using quantum computers*, *Molecular Physics* **109** (2011) 735.
- [4] J. Lee, D.W. Berry, C. Gidney, W.J. Huggins, J.R. McClean, N. Wiebe et al., *Even more efficient quantum computations of chemistry through tensor hypercontraction*, *PRX Quantum* **2** (2021) 030305.
- [5] V. von Burg, G.H. Low, T. Häner, D.S. Steiger, M. Reiher, M. Roetteler et al., *Quantum computing enhanced computational catalysis*, *Physical Review Research* **3** (2021) 033055.
- [6] S.P. Jordan, K.S. Lee and J. Preskill, *Quantum algorithms for quantum field theories*, *Science* **336** (2012) 1130.
- [7] A.F. Shaw, P. Lougovski, J.R. Stryker and N. Wiebe, *Quantum algorithms for simulating the lattice schwinger model*, *Quantum* **4** (2020) 306.
- [8] N. Klco, M.J. Savage and J.R. Stryker, *Su(2) non-abelian gauge field theory in one dimension on digital quantum computers*, *Physical Review D* **101** (2020) 074512.
- [9] A.M. Childs and N. Wiebe, *Hamiltonian simulation using linear combinations of unitary operations*, *Quantum Info. Comput.* **12** (2012) 901–924.
- [10] G.H. Low, V. Kliuchnikov and N. Wiebe, *Well-conditioned multiproduct hamiltonian simulation*, *arXiv:1907.11679* (2019) .

- [11] D.W. Berry, A.M. Childs, R. Cleve, R. Kothari and R.D. Somma, *Simulating hamiltonian dynamics with a truncated taylor series*, *Physical review letters* **114** (2015) 090502.
- [12] G.H. Low and N. Wiebe, *Hamiltonian simulation in the interaction picture*, *arXiv:1805.00675* (2018) .
- [13] M. Kieferová, A. Scherer and D.W. Berry, *Simulating the dynamics of time-dependent hamiltonians with a truncated dyson series*, *Physical Review A* **99** (2019) 042314.
- [14] G.H. Low and I.L. Chuang, *Hamiltonian Simulation by Qubitization*, *Quantum* **3** (2019) 163.
- [15] R. Babbush, C. Gidney, D.W. Berry, N. Wiebe, J. McClean, A. Paler et al., *Encoding electronic spectra in quantum circuits with linear t complexity*, *Physical Review X* **8** (2018) 041015.
- [16] D.W. Berry, G. Ahokas, R. Cleve and B.C. Sanders, *Efficient quantum algorithms for simulating sparse hamiltonians*, *Communications in Mathematical Physics* **270** (2006) 359–371.
- [17] N. Wiebe, D.W. Berry, P. Høyer and B.C. Sanders, *Simulating quantum dynamics on a quantum computer*, *Journal of Physics A: Mathematical and Theoretical* **44** (2011) 445308.
- [18] A.M. Childs, Y. Su, M.C. Tran, N. Wiebe and S. Zhu, *Theory of trotter error with commutator scaling*, *Physical Review X* **11** (2021) 011020.
- [19] J. Haah, M.B. Hastings, R. Kothari and G.H. Low, *Quantum algorithm for simulating real time evolution of lattice hamiltonians*, *SIAM Journal on Computing* (2021) FOCS18.
- [20] M. Hagan and N. Wiebe, *Composite quantum simulations*, *arXiv:2206.06409* (2022) .
- [21] G.H. Low, Y. Su, Y. Tong and M.C. Tran, *On the complexity of implementing trotter steps*, *arXiv:2211.09133* (2022) .
- [22] G.H. Low and I.L. Chuang, *Optimal hamiltonian simulation by quantum signal processing*, *Physical Review Letters* **118** (2017) .
- [23] S. Endo, Q. Zhao, Y. Li, S. Benjamin and X. Yuan, *Mitigating algorithmic errors in a hamiltonian simulation*, *Phys. Rev. A* **99** (2019) 012334.
- [24] A.C. Vazquez, R. Hiptmair and S. Woerner, *Enhancing the quantum linear systems algorithm using richardson extrapolation*, *ACM Transactions on Quantum Computing* **3** (2022) .

- [25] A.C. Vazquez, D.J. Egger, D. Ochsner and S. Woerner, *Well-conditioned multi-product formulas for hardware-friendly hamiltonian simulation*, *Quantum* **7** (2023) 1067.
- [26] M. Suzuki, *General theory of fractal path integrals with applications to many-body theories and statistical physics*, *Journal of Mathematical Physics* **32** (1991) 400.
- [27] A. Gilyén, Y. Su, G.H. Low and N. Wiebe, *Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics*, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pp. 193–204, 2019, DOI.
- [28] C. Yi and E. Crosson, *Spectral analysis of product formulas for quantum simulation*, *npj Quantum Information* **8** (2022) 37.
- [29] A. Quarteroni, R. Sacco and F. Saleri, *Numerical mathematics*, vol. 37, Springer Science & Business Media (2010), [10.1007/b98885](https://doi.org/10.1007/b98885).
- [30] F. Piazzon and M. Vianello, *Stability inequalities for lebesgue constants via markov-like inequalities*, *Dolomites Research Notes on Approximation* **11** (2018) .
- [31] A.P. de Camargo, *On the numerical stability of newton’s formula for lagrange interpolation*, *Journal of Computational and Applied Mathematics* **365** (2020) 112369.
- [32] L. Trefethen, *Six myths of polynomial interpolation and quadrature*, .
- [33] W. Gautschi, *How (un)stable are vandermonde systems? asymptotic and computational analysis*, in *Lecture Notes in Pure and Applied Mathematics*, pp. 193–210, Marcel Dekker, Inc, 1990.
- [34] N.J. Higham, *The numerical stability of barycentric lagrange interpolation*, *IMA Journal of Numerical Analysis* **24** (2004) 547.
- [35] J.C. Mason and D.C. Handscomb, *Chebyshev polynomials*, CRC press (2002), [10.1201/9781420036114](https://doi.org/10.1201/9781420036114).
- [36] G. Rendon, T. Izubuchi and Y. Kikuchi, *Effects of cosine tapering window on quantum phase estimation*, *Physical Review D* **106** (2022) 034503.
- [37] L.N. Trefethen, *Approximation Theory and Approximation Practice, Extended Edition*, SIAM (2019), [10.1137/1.9781611975949](https://doi.org/10.1137/1.9781611975949).
- [38] F.L. Bauer and C.T. Fike, *Norms and exclusion theorems*, *Numer. Math.* **2** (1960) 137–141.
- [39] S. Blanes, F. Casas, J.-A. Oteo and J. Ros, *The magnus expansion and some of its applications*, *Physics reports* **470** (2009) 151.

- [40] N. Klco and M.J. Savage, *Minimally entangled state preparation of localized wave functions on quantum computers*, *Physical Review A* **102** (2020) .
- [41] J.J. García-Ripoll, *Quantum-inspired algorithms for multivariate analysis: from interpolation to partial differential equations*, *Quantum* **5** (2021) 431.
- [42] W. Górecki, R. Demkowicz-Dobrzański, H.M. Wiseman and D.W. Berry, *π -corrected heisenberg limit*, *Physical review letters* **124** (2020) 030501.
- [43] D. Grinko, J. Gacon, C. Zoufal and S. Woerner, *Iterative quantum amplitude estimation*, *npj Quantum Information* **7** (2021) 52 [[1912.05559](#)].
- [44] N. Wiebe, D. Berry, P. Høyer and B.C. Sanders, *Higher order decompositions of ordered operator exponentials*, *Journal of Physics A: Mathematical and Theoretical* **43** (2010) 065203.
- [45] R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge university press (2012), [10.1017/CBO9780511810817](#).
- [46] M. Chiani, D. Dardari and M.K. Simon, *New exponential bounds and approximations for the computation of error probability in fading channels*, *IEEE Transactions on Wireless Communications* **2** (2003) 840.
- [47] J.M. Borwein and P.B. Borwein, *Pi and the AGM: a study in the analytic number theory and computational complexity*, Wiley-Interscience (1987).
- [48] B.L. Higgins, D.W. Berry, S.D. Bartlett, H.M. Wiseman and G.J. Pryde, *Entanglement-free Heisenberg-limited phase estimation*, *Nature* **450** (2007) 393.
- [49] R.B. Griffiths and C.-S. Niu, *Semiclassical Fourier Transform for Quantum Computation*, *Physical Review Letters* **76** (1996) 3228.
- [50] A.Y. Kitaev, *Quantum measurements and the abelian stabilizer problem*, [quant-ph/9511026](#) (1995) .
- [51] D.S. Abrams and S. Lloyd, *Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors*, *Physical Review Letters* **83** (1999) 5162.
- [52] J. Watkins, N. Wiebe, A. Roggero and D. Lee, *Time-dependent hamiltonian simulation using discrete clock constructions*, [arXiv:2203.11353](#) (2022) .
- [53] T.D. Ahle, *Sharp and simple bounds for the raw moments of the binomial and poisson distributions*, *Statistics & Probability Letters* **182** (2022) 109306.
- [54] T. Rivlin, *Chebyshev Polynomials*, Dover Books on Mathematics, Dover Publications (2020).

A Proof of Lemma 1: derivatives of effective Hamiltonian

Proof. Recall the definition of the effective Hamiltonian

$$\tilde{H}_s := \frac{i}{st} \log S_{2k}(st) \quad (\text{A.1})$$

for $s \in \mathbb{R} \setminus \{0\}$, with $\tilde{H}_0 := \lim_{s \rightarrow 0} \tilde{H}_s = H$. We will understand $\log S_{2k}(st)$ through a power series expansion about the identity.

$$\log S_{2k}(st) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j+1} (S_{2k}(st) - \mathbb{1})^{j+1} \quad (\text{A.2})$$

This series converges precisely when

$$\|S_{2k}(st) - \mathbb{1}\| \leq 1. \quad (\text{A.3})$$

Using the fundamental theorem of calculus, we can derive a suitable condition for convergence as a neighborhood about $s = 0$. The condition above implies

$$\left\| \int_0^{st} S_{2k}(x) dx \right\| \leq 1 \quad (\text{A.4})$$

which is satisfied provided

$$|st| \max_{x \in [0, st]} \left\| \frac{d}{dx} S_{2k}(x) \right\| \leq 1. \quad (\text{A.5})$$

Writing out $S_{2k}(x) = \prod_{l=1}^{N_k} \exp(-iH_{j_l} \tau_l x)$ where H_{j_l} is some Hamiltonian piece H_j indexed by l , the derivative can be upper bounded as

$$\begin{aligned} \max_{x \in [0, st]} \left\| \frac{d}{dx} S_{2k}(x) \right\| &\leq \sum_{l=1}^{N_k} \|H_{j_l}\| |\tau_l| \\ &\leq \max_j \|H_j\| \|\tau\|_1 \end{aligned} \quad (\text{A.6})$$

where $\tau = (\tau_l)_{l=1}^{N_k}$ is the vector of ST coefficients, and in going to the second line we used a Hölder inequality. We have $\|\tau\|_1 \leq N_k \max_l |\tau_l|$, and from Appendix A of [44] we have

$$\max_l |\tau_l| \leq 2k/3^k. \quad (\text{A.7})$$

Thus, the requirement for convergence of the logarithm becomes

$$\frac{4}{3} k (5/3)^{k-1} m |st| \max_j \|H_j\| \leq 1 \quad (\text{A.8})$$

where we used the expression $N_k = (2m)5^{k-1}$ for the number of ST exponentials.

We now assume s is within the symmetric interval defined by (A.8), such that (A.2) is convergent. Since $\log S_{2k}(0) = 0$, $s = 0$ is a zero of order at least one. We want to absorb the diverging $1/s$ term and better understanding the leading dependence in s . To facilitate this, we write

$$\tilde{H}_s = -\frac{1}{it} \sum_{j=0}^{\infty} \frac{(-1)^j}{j+1} s^j \Delta S_{2k}(st)^{j+1} \quad (\text{A.9})$$

where we defined

$$\Delta S_{2k}(st) := \frac{S_{2k}(st) - \mathbb{1}}{s}. \quad (\text{A.10})$$

Note that ΔS_{2k} is analytic in s , and is a finite difference around $s = 0$, such that

$$\lim_{s \rightarrow 0} \Delta S_{2k}(st) = -iHt. \quad (\text{A.11})$$

Through the series expansion (A.9) we may bound derivatives of \tilde{H}_s via bounds on derivatives of ΔS_{2k} . We first obtain a power series of ΔS_{2k} by Taylor expanding every term in the product formula S_{2k} . Regrouping in powers of st , the result is

$$\Delta S_{2k}(st) = \sum_{j=1}^{\infty} \frac{s^{j-1}(-it)^j}{j!} \sum_J \binom{j}{j_1 \dots j_{N_k}} \prod_{l=1}^{N_k} (H_l \tau_l)^{j_l} \quad (\text{A.12})$$

where the parenthetical symbol is the multinomial coefficient, and the sum \sum_J is over all values of $J = (j_1, \dots, j_{N_k})$ such that $\sum_k j_k = j$. The derivatives with respect to s are now easy to compute. Using the fact that

$$\partial_s^n s^{j-1} = \frac{(j-1)!}{(j-1-n)!} s^{j-n-1} \quad (\text{A.13})$$

for $j > n$ (and zero otherwise), we have

$$\begin{aligned} \partial_s^n \Delta S_{2k}(st) &= \sum_{j=n+1}^{\infty} \frac{s^{j-n-1}(-it)^j}{j!} \frac{(j-1)!}{(j-1-n)!} \sum_J \binom{j}{j_1 \dots j_{N_k}} \prod_{l=1}^{N_k} (H_l \tau_l)^{j_l} \\ \|\partial_s^n \Delta S_{2k}(st)\| &\leq \sum_{j=n+1}^{\infty} \frac{t^j}{(j-n-1)!} s^{j-n-1} (\tau_{\max} N_k \Lambda)^j \end{aligned} \quad (\text{A.14})$$

where $\Lambda := \max_j \|H_j\|$ and $\tau_{\max} = \max_l |\tau_l|$. Factoring out powers of $n+1$ and reindexing, we are left with the following bound on derivatives of ΔS_{2k} .

$$\|\partial_s^n \Delta S_{2k}(st)\| \leq (\tau_{\max} N_k \Lambda t)^{n+1} e^{s\tau_{\max} N_k \Lambda t} \quad (\text{A.15})$$

This expression is quite elegant; it is as if we were taking $n+1$ derivatives of the exponential e^{cs} with

$$\begin{aligned} c &:= \tau_{\max} N_k \Lambda t \\ &\leq k(5/3)^k m \Lambda t \end{aligned} \quad (\text{A.16})$$

Factors of c will occur frequently in what follows, so we find it convenient to adopt this symbol as shorthand.

We return to bounding the derivatives of powers of $\Delta S_{2k}(st)$ as in equation (A.9).

$$\partial_s^n \left[\Delta S_{2k}(st)^{j+1} \right] \quad (\text{A.17})$$

We reduce this to the previous case by performing a multinomial expansion.

$$\partial_s^n \Delta S_{2k}(st)^{j+1} = \sum_N \binom{n}{n_0 \dots n_j} \prod_{l=0}^j \partial_s^{n_l} \Delta S_{2k}(st) \quad (\text{A.18})$$

As usual, the capital letter N denotes the set of all nonnegative indices n_0, \dots, n_j summing to n . Applying the triangle inequality and submultiplicativity, and employing the bound (A.15),

$$\begin{aligned} \left\| \partial_s^n \Delta S_{2k}(st)^{j+1} \right\| &\leq \sum_N \binom{n}{n_0 \dots n_j} \prod_{l=0}^j \left\| \partial_s^{n_l} \Delta S_{2k}(st) \right\| \\ &\leq \sum_N \binom{n}{n_0 \dots n_j} \prod_{l=0}^j c^{n_l+1} e^{cs} \\ &= e^{(j+1)cs} c^{n+j+1} \sum_N \binom{n}{n_0 \dots n_j}, \end{aligned} \quad (\text{A.19})$$

where we've used the sum property of the n_l where appropriate. The remaining sum over the multinomial coefficient is given by $(j+1)^n$. Hence,

$$\left\| \partial_s^n \Delta S_{2k}(st)^{j+1} \right\| \leq ((j+1)c)^n (ce^{cs})^{j+1} \quad (\text{A.20})$$

Notice that, when $j=0$, this is consistent with equation (A.15).

With result (A.20) in hand, we return to the power series (A.9). Differentiating term by term

$$\partial_s^n \tilde{H}_s = -\frac{1}{it} \sum_{j=0}^{\infty} \frac{(-1)^j}{j+1} \partial_s^n \left(s^j \Delta S_{2k}(st)^{j+1} \right) \quad (\text{A.21})$$

and performing a binomial expansion for each term

$$\partial_s^n \left(s^j \Delta S_{2k}(st)^{j+1} \right) = \sum_{q=0}^n \binom{n}{q} \left(\partial_s^q s^j \right) \left(\partial_s^{n-q} \Delta S_{2k}(st)^{j+1} \right) \quad (\text{A.22})$$

will allow us to apply our previous results. It will be helpful to consider two cases separately: $j \leq n$ and $j > n$. These regimes are somewhat qualitatively different, since the derivatives of s^j may or may not vanish depending on the number of derivatives. Focusing on the case $j \leq n$, we have

$$\partial_s^n \left(s^j \Delta S_{2k}(st)^{j+1} \right) = \sum_{q=0}^j \binom{n}{q} \frac{j!}{(j-q)!} s^{j-q} \left(\partial_s^{n-q} \Delta S_{2k}(st)^{j+1} \right). \quad (\text{A.23})$$

Note that the sum runs only to j , not n . Taking a triangle inequality upper bound using (A.20), we may upper bound (A.23) as

$$\begin{aligned} & \sum_{q=0}^j \binom{n}{q} \frac{j!}{(j-q)!} s^{j-q} ((j+1)c)^{n-q} (ce^{cs})^{j+1} \\ &= (ce^{cs})^{j+1} \sum_{q=0}^j \binom{j}{q} \frac{n!}{(n-q)!} s^{j-q} ((j+1)c)^{n-q} \end{aligned} \quad (\text{A.24})$$

where we have factored out terms not involving q from the sum, and manipulated the factorials for reasons which will be seen presently. Taking the upper bound $n!/(n-q)! < n^q$, and factoring out $n-j$ powers of $(j+1)c$, we may upper bound the above expression by

$$\begin{aligned} & (ce^{cs})^{j+1} ((j+1)c)^{n-j} \sum_{q=0}^j \binom{j}{q} n^q ((j+1)cs)^{j-q} \\ &= (ce^{cs})^{j+1} ((j+1)c)^{n-j} (n + (j+1)cs)^j \end{aligned} \quad (\text{A.25})$$

Thus, with some minor polishing, we may express the bound on (A.22) for $j \leq n$ as

$$\left\| \partial_s^n \left(s^j \Delta S_{2k}(st)^{j+1} \right) \right\| \leq e^{(j+1)cs} c^{n+1} (j+1)^n \left(\frac{n}{j+1} + cs \right)^j. \quad (\text{A.26})$$

Now let's move on to the $j > n$ case. Here, there are not enough derivatives to kill off the s^j term, so the binomial sum in (A.23) will run from $q = 0$ to n .

$$\partial_s^n \left(s^j \Delta S_{2k}(st)^{j+1} \right) = \sum_{q=0}^n \binom{n}{q} \frac{j!}{(j-q)!} s^{j-q} \left(\partial_s^{n-q} \Delta S_{2k}(st)^{j+1} \right) \quad (\text{A.27})$$

Similar to before, we use the bound (A.20), to obtain

$$\begin{aligned} \left\| \partial_s^n \left(s^j \Delta S_{2k}(st)^{j+1} \right) \right\| &\leq \sum_{q=0}^n \binom{n}{q} \frac{j!}{(j-q)!} s^{j-q} ((j+1)c)^{n-q} (ce^{cs})^{j+1} \\ &= (ce^{cs})^{j+1} s^{j-n} \sum_{q=0}^n \binom{n}{q} \frac{j!}{(j-q)!} ((j+1)cs)^{n-q}. \end{aligned} \quad (\text{A.28})$$

Taking $j!/(j-q)! < j^q$, a simpler upper bound is given by

$$(ce^{cs})^{j+1} s^{j-n} \sum_{q=0}^n \binom{n}{q} j^q ((j+1)cs)^{n-q} = (ce^{cs})^{j+1} s^{j-n} (j + (j+1)cs)^n. \quad (\text{A.29})$$

With some minor rearrangements, this gives the following upper bound for $j > n$.

$$\left\| \partial_s^n \left(s^j \Delta S_{2k}(st)^{j+1} \right) \right\| \leq e^{(j+1)cs} c^{n+1} (j+1)^n (cs)^{j-n} \left(\frac{j}{j+1} + cs \right)^n \quad (\text{A.30})$$

With the bounds (A.26) and (A.30), we can return to bounding $\partial_s^n \tilde{H}_s$. Still separating the two cases $j \leq n$ and $j > n$, we can write

$$\begin{aligned} \|\partial_s^n \tilde{H}_s\| t &\leq \sum_{j=0}^n \frac{1}{j+1} \|\partial_s^n (s^j \Delta S_{2k}(st)^{j+1})\| + \sum_{j=n+1}^{\infty} \frac{1}{j+1} \|\partial_s^n (s^j \Delta S_{2k}(st)^{j+1})\| \\ &= B_l + B_h \end{aligned} \quad (\text{A.31})$$

where B_l and B_h refer to bounds on the “low” and “high” parts of the series. Employing the bounds from equations (A.26) and (A.30), we have

$$\begin{aligned} B_l &\leq \sum_{j=0}^n \frac{1}{j+1} e^{(j+1)cs} c^{n+1} (j+1)^n \left(\frac{n}{j+1} + cs \right)^j \\ &= c^{n+1} \sum_{j=0}^n e^{(j+1)cs} (j+1)^{n-1} \left(cs + \frac{n}{j+1} \right)^j \end{aligned} \quad (\text{A.32})$$

and

$$\begin{aligned} B_h &\leq \sum_{j=n+1}^{\infty} \frac{1}{j+1} e^{(j+1)cs} c^{n+1} (j+1)^n (cs)^{j-n} \left(\frac{j}{j+1} + cs \right)^n \\ &\leq c^{n+1} \sum_{j=n+1}^{\infty} e^{(j+1)cs} (j+1)^{n-1} (cs)^{j-n} (1+cs)^n \\ &= c^{n+1} (1+cs)^n \sum_{j=n+1}^{\infty} e^{(j+1)cs} (j+1)^{n-1} (cs)^{j-n}. \end{aligned} \quad (\text{A.33})$$

Let’s begin by simplifying the bound on B_l . We will at this point make the assumption that s is sufficiently small such that $cs < 1$. This will necessarily factor into the cost later. This simplification yields

$$\begin{aligned} B_l &\leq c^{n+1} \sum_{j=0}^n e^{j+1} (j+1)^{n-1} \left(1 + \frac{n}{j+1} \right)^j \\ &\leq c^{n+1} \sum_{j=0}^n e^{j+1} (j+1)^{n-1} e^n \\ &\leq \frac{1}{e} (e^2 c)^{n+1} \sum_{j=1}^{n+1} j^{n-1} \end{aligned} \quad (\text{A.34})$$

The remaining sum can be bounded by $(n+1)^n$, hence,

$$B_l \leq \frac{(e^2(n+1)c)^{n+1}}{e(n+1)} \leq (e^2 c)^{n+1} n^n, \quad (\text{A.35})$$

where the definition that $0^0 = 1$ handles the edge case. Let’s turn our attention to B_h . We will start by reindexing so that the series begins at $j = 0$ in (A.33).

$$B_h \leq c^{n+1} (1+cs)^n \sum_{j=0}^{\infty} e^{(j+n+2)cs} (j+n+2)^{n-1} (cs)^{j+1} \quad (\text{A.36})$$

$$= (ce^{cs})^{n+1}(1+cs)^n \sum_{j=0}^{\infty} (cse^{cs})^{j+1} (j+n+2)^{n-1} \quad (\text{A.37})$$

The series converges if and only if

$$cse^{cs} < 1. \quad (\text{A.38})$$

This condition is slightly stronger than the condition (A.8) that we need for convergence of the logarithm (A.2), and is equivalent to $cs < W(1) \approx 0.567$, where W is the principal branch of the Lambert W function. Returning to (A.36), we have the bound

$$(j+n+2)^{n-1} = (n+2)^{n-1} \left(1 + \frac{j}{n+2}\right)^{n-1} \leq (n+2)^{n-1} e^j. \quad (\text{A.39})$$

Thus, we have

$$\begin{aligned} B_h &\leq (ce^{cs})^{n+1}(1+cs)^n (n+2)^{n-1} \sum_{j=0}^{\infty} (ecse^{cs})^j \\ &= (ce^{cs})^{n+1}(1+cs)^n (n+2)^{n-1} \frac{1}{1-ecse^{cs}}. \end{aligned} \quad (\text{A.40})$$

To be concrete, let's take $ecse^{cs} < 1/2$, which is implied by $cs < \pi/20$. Coupled with the inequality in (A.16), this condition can be met provided that

$$k(5/3)^k m \Lambda s t \leq \pi/20, \quad (\text{A.41})$$

which is exactly the assumption of Lemma 1. This allows us to upper bound B_h further as

$$B_h \leq 2e^{\pi/20} (cs)^{n+1} (3e^{\pi/20}/2)^n (n+2)^{n-1} \leq 4(cs)^{n+1} (9/5)^n (n+2)^{n-1}. \quad (\text{A.42})$$

Since $(n+2)^{n-1} \leq e^2 n^n / 2$ (using $0^0 := 1$ for the edge case $n = 0$), we have

$$B_h \leq 2e^2 (cs)^{n+1} (9/5)^n n^n. \quad (\text{A.43})$$

Altogether, using $s \leq 1$

$$\begin{aligned} \left\| \partial_s^n \tilde{H}_s \right\| t &\leq n^n (e^2 c)^{n+1} \left(1 + 2(9/5e^2)^n\right) \\ &\leq 2n^n (e^2 c)^{n+1} \end{aligned} \quad (\text{A.44})$$

The final result then follows from substituting for c and noting that the duration of each time step is at most $2k/3^{k-1}$ using the results of [17]. \square

B Proof of Theorem 7

In order to use the result of Lemma 1 in the proof of Theorem 7, we need to re-express the derivatives of the eigenvalues in terms of the effective Hamiltonian \tilde{H}_t (we use t rather than s due to the extrapolation parameter taking the role of time). The following lemma provides such a bound on the derivatives under the assumption that the Hamiltonian is gapped.

Lemma 20. *Let $\tilde{H}_t : \mathbb{R} \mapsto \mathbb{C}^{2^n \times 2^n}$ be a Hamiltonian on n qubits with p continuous derivatives. Suppose \tilde{H}_t has a minimum eigenvalue gap*

$$\gamma(t) := \min_{k \neq j} |\lambda_k(t) - \lambda_j(t)|$$

where $\lambda_j(t)$ and $\lambda_k(t)$ are eigenvalues of \tilde{H}_t . Define

$$\Lambda_\gamma := \max_p \sup_t \left(\|\partial_t^p \tilde{H}_t\|^{1/(p+1)}, \left(\frac{\|\partial_t^p \tilde{H}_t\|}{\gamma(t)} \right)^{1/p} \right),$$

which has dimensions of $1/t$. Then for all eigenvalues $\lambda_k(t)$

$$|\partial_t^p \lambda_k(t)| \leq 8^{p-1} (p-1)! \Lambda_\gamma^{p+1}.$$

Proof. The essence of this proof is to examine the derivatives of $\lambda_k(t)$ by expressing them in terms of derivatives of \tilde{H}_t via perturbation theory. Using the bound Lemma 1 on \tilde{H}_t lets us bound the eigenvalues in turn. Let $|k(t)\rangle$ be the eigenvector for eigenvalue $\lambda_k(t)$ and assume that the eigenvectors' phases are chosen such that $\langle k(t) | \dot{k}(t) \rangle = 0$ for all t . Such a choice is always possible. Standard results from perturbation theory [45] show that

$$\begin{aligned} \partial_t \lambda_k(t) &= \langle k(t) | (\partial_t \tilde{H}_t) | k(t) \rangle \\ \partial_t |k(t)\rangle &= \sum_{j \neq k} |j(t)\rangle \frac{\langle j(t) | \partial_t \tilde{H}_t | k(t) \rangle}{\lambda_k(t) - \lambda_j(t)}. \end{aligned} \quad (\text{B.1})$$

Further, taking the Euclidean norm of the 2nd equation,

$$\|\partial_t |k(t)\rangle\|^2 \leq \sum_{j \neq k} \left| \frac{\langle j(t) | \partial_t \tilde{H}_t | k(t) \rangle}{\lambda_k(t) - \lambda_j(t)} \right|^2 \leq \frac{\|\partial_t \tilde{H}_t\|^2}{\gamma^2(t)}. \quad (\text{B.2})$$

where $\gamma(t)$ is the minimum eigenvalue gap of the effective Hamiltonian at time t .

We now consider second- and higher-order derivatives of $\lambda_k(t)$. By repeatedly taking derivatives of $\lambda_k(t)$ in the first line of (B.1), we obtain an expression of the general form

$$\partial_t^q \lambda_k(t) = \sum_{\ell=1}^q \sum_{x,y} c_{\vec{x},\vec{y},\ell} \langle x_1(t) | \partial_t^{y_\mu} \tilde{H}_t | x_2(t) \rangle \prod_{\mu=2}^{\ell} \frac{\langle x_{2\mu-1}(t) | \partial_t^{y_\mu} \tilde{H}_t | x_{2\mu}(t) \rangle}{(\lambda_{x_{2\mu}}(t) - \lambda_{x_{2\mu-1}}(t))}. \quad (\text{B.3})$$

Note that there is a pattern to this series. If there are m $H^{(p)}$ (for $p \geq 0$) present in the expansion there must be $m - 1$ powers of the gaps present in the terms. Before going into more detail let us define the “degree” of a term to be the number of differentiable factors (eigenvectors, Hamiltonians or inverse gaps) present in a product. For example $\langle x(t) | \partial_t \tilde{H}_t | x(t) \rangle$ is degree 3.

We have that if A is a term of degree $\deg(A)$. Then it follows from (B.1) that

$$\deg(\partial_t A) \leq \deg(A) + 4, \quad (\text{B.4})$$

since from the product rule the derivative of the product of the factors is the sum of the distributed sum of the derivatives of the factors and that if $|\lambda_k\rangle$ or $(\lambda_j - \lambda_k)^{-1}$ are differentiated then the degree increases by at most 4. While the derivative of the effective Hamiltonian does not increase the degree, the increase in the degree is still at most 4 in this case thus in the worst case scenario the degree is increased by 4. As the degree $\deg(\partial_t \lambda_k(t)) = 3 \leq 4$, it is straight forward to see that

$$\deg(\partial_t^q \lambda_k(t)) = 4q \quad (\text{B.5})$$

Next, let us assume that B is a sum of $N_{\text{terms}}(B)$ products of the above factors. We then have from (B.1) that

$$N_{\text{terms}}(\partial_t B) \leq 2N_{\text{terms}}(B)\deg(B). \quad (\text{B.6})$$

Next we will show that the number of terms present in $\partial_t^q \lambda_k(t)$ obeys

$$N_{\text{terms}}(\partial_t^q \lambda_k(t)) \leq 8^{q-1}(q-1)!. \quad (\text{B.7})$$

We prove this by induction. The number of terms present when $q = 1$ is 1. This demonstrates the base case. Assume that for some value q' that the induction hypothesis holds. Then from (B.5) and (B.6) that

$$N_{\text{terms}}(\partial_t^{q'+1} \lambda_k(t)) \leq 2(8^{q'-1}(q'-1)!(4q')) = 8^{q'} q'! \quad (\text{B.8})$$

which proves the relation by induction.

We then see that if any factor is differentiated ν times then its norm is bounded above by $\Lambda_\gamma^{\nu+1}$ where

$$\Lambda_\gamma := \max_p \left(\max \left\{ \|\partial_t^p \tilde{H}_t\|^{1/(p+1)}, \left(\frac{\|\partial_t^p \tilde{H}_t\|}{\gamma(t)} \right)^{1/p} \right\} \right). \quad (\text{B.9})$$

We therefore have that since the maximum value of the derivative is the number of terms multiplied by the maximum norm of the product of all the factors that

$$|\partial_t^q \lambda_k(t)| \leq 8^{q-1}(q-1)!\Lambda_\gamma^{q+1}. \quad (\text{B.10})$$

□

Proof of Theorem 7. First, from the Baker Campbell Hausdorff formula,

$$S_{2k}(t) = \exp\left(-iHt + O(t^{2k+1})\right)$$

and thus $\log(S_{2k}(t))/t \in O(1)$ as $t \rightarrow 0$. Thus, an eigenvalue $E_{\text{eff}}(t)$ exists for \tilde{H}_t for all $t > 0$, demonstrating the first claim.

Next, since \tilde{H}_t has p continuous derivatives, $\partial_t^p \lambda_k(t)$ exists and is bounded above by $8^{p-1}(p-1)!\Lambda_\gamma^{p+1}$ because

$$\partial_t \tilde{H}_t = \frac{1}{t} \partial_s \tilde{H}_{(st)}. \quad (\text{B.11})$$

The value of Λ_γ can then be bounded, using Lemma 1, by

$$\begin{aligned} \Lambda_\gamma &\leq \left(2t^{-n-1}n^n(e^2k(5/3)^{k-1}m \max_{l \in [1,m]} \|H_l\|t)^{n+1}\right)^{1/(n+1)} \\ &\quad + \max_p \left(\frac{2t^{-p-1}p^p(e^2k(5/3)^{k-1}m \max_{l \in [1,m]} \|H_l\|t)^{p+1}}{\gamma}\right)^{1/p} \\ &\leq 4nke^2(5/3)^{k-1}m \max_{l \in [1,m]} \|H_l\| \left(1 + \max_p \left(\frac{p}{n}\right) \left(\frac{e^2k(5/3)^{k-1}m \max_{l \in [1,m]} \|H_l\|}{\gamma}\right)^{1/p}\right) \\ &:= 4nke^2(5/3)^{k-1}m \max_{l \in [1,m]} \|H_l\| (1 + \Gamma). \end{aligned} \quad (\text{B.12})$$

This implies that

$$|\partial_t^q \lambda_k(t)| \leq 8^{q-1}(q-1)!\Lambda_\gamma^{q+1} \leq \frac{q!q^q}{64} \left(32ke^2(5/3)^{k-1}m \max_{l \in [1,m]} \|H_l\| (1 + \Gamma)\right)^{q+1}. \quad (\text{B.13})$$

For all $t \in [-a, a]$, we have from Lemma 2 that the interpolation error satisfies the following upper bound.

$$|\lambda_k(0) - P(0)| \leq \frac{n!}{64} \left(32ke^2(5/3)^{k-1}m \max_{l \in [1,m]} \|H_l\| (1 + \Gamma)\right)^{n+1} \left(\frac{a}{2}\right)^n \quad (\text{B.14})$$

We wish for this error to be at most ϵ_{int} . Using the inequality above, we can solve for a value of a such that the error is within ϵ_{int} for all $t \in [-a, a]$. Doing so, we find the following value of a .

$$a = \frac{2(64\epsilon_{\text{int}})^{1/n}}{(n!)^{1/n}} \left(\frac{1}{32ke^2(5/3)^{k-1}m \max_{l \in [1,m]} \|H_l\| (1 + \Gamma)}\right)^{1+1/n} \quad (\text{B.15})$$

We now count the number of exponentials needed for the degree n polynomial interpolation (where n is even) using (B.15). Using the fact that the error for time $-t$ is equal to the error for time t , that the j^{th} Chebyshev node s_j scales like a/j , and

that the cost of phase estimation within error ϵ_{PE} with failure probability δ_{PE} scales as $O(\log(1/\delta_{\text{PE}})/\epsilon_{\text{PE}})$, the number of exponentials scales as

$$\begin{aligned}
N_{\text{exp}} &\in O\left(\sum_{j=1}^{n/2} \frac{m5^{k-1}}{ja}\right) \\
&= O\left(\frac{n \log(n) m 5^{k-1} \log(1/\delta_{\text{PE}})}{a \epsilon_{\text{PE}}}\right) \\
&= O\left(\frac{n \log(n) \log(1/\delta_{\text{PE}}) (n!)^{1/n} m^{2+1/n} (5/3)^{2k+k/n} (\max \|H_i\| (1+\Gamma) k)^{1+1/n}}{\epsilon_{\text{int}}^{1/n} \epsilon_{\text{PE}}}\right) \\
&= \tilde{O}\left(\frac{n^2 m^{2+1/n} \log(1/\delta_{\text{PE}}) 5^k (5/3)^{k+k/n} (\max \|H_i\| (1+\Gamma) k)^{1+1/n}}{\epsilon_{\text{int}}^{1/n} \epsilon_{\text{PE}}}\right). \tag{B.16}
\end{aligned}$$

There are two competing tendencies in the cost. The number of operator exponentials increases polynomially with n , but the scaling with the error tolerance improves exponentially with n . Setting these two equal to each other to estimate the optimal scaling yields

$$n^2 = \left(\frac{mk(5/3)^k \max_i \|H_i\| (1+\Gamma)}{\epsilon_{\text{int}}}\right)^{1/n}. \tag{B.17}$$

Under the assumption that $\epsilon_{\text{int}} \leq 5/3$, the solution is of the form

$$\begin{aligned}
n &= \frac{\ln\left(mk(5/3)^k \max_i \|H_i\| (1+\Gamma) \epsilon_{\text{int}}^{-1}\right)}{2 \text{LambertW}\left(\ln\left(mk(5/3)^k \max_i \|H_i\| (1+\Gamma) \epsilon_{\text{int}}^{-1}\right) / 2\right)} \\
&\in O\left(\log\left(mk(5/3)^k \max_i \|H_i\| (1+\Gamma) \epsilon_{\text{int}}^{-1}\right)\right). \tag{B.18}
\end{aligned}$$

Finally, we have from Theorem 5 that the error in the interpolated eigenvalue is within ϵ provided that the error in the data satisfies

$$\left(\frac{2}{\pi} \log(n+1) + 1\right) \max_i |\lambda_k(s_i) - \tilde{\lambda}_k(s_i)| \leq \left(\frac{2}{\pi} \log(n+1) + 1\right) \epsilon_{\text{PE}} \leq \epsilon. \tag{B.19}$$

Taking ϵ_{PE} to satisfy this, the number of exponentials scales via equation (B.16) as

$$N_{\text{exp}} \in \tilde{O}\left(\frac{m^2 k (25/3)^k \max \|H_i\| (1+\Gamma)}{\epsilon}\right), \tag{B.20}$$

where we've also implemented (B.17). By the union bound, the total probability of failure is at most $1/3$. □

C Proof of Theorem 16

Per the statement of the theorem, we would like to estimate the error between the Discrete Fourier Transform (DFT) and corresponding samples of the Fourier transform of the continuous Gaussian distribution. We split this analysis in three parts: calculating time-domain truncation errors, estimating frequency-domain truncation errors (non-zero sampling rate), and finally calculate the error coming from renormalization. Figure 6 gives an schematic representation of the analysis, including the resulting big- O bounds.

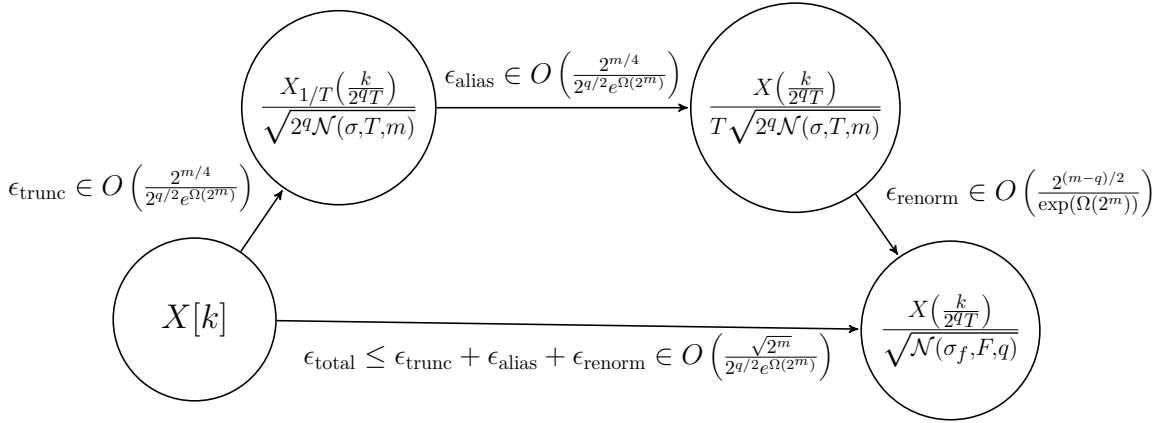


Figure 6: Schematic representation of the error estimation between the DFT of $x[n]$, $X[k]$, and the re-normalized samples of the Fourier transform $X\left(\frac{k}{2^q T}\right)$. At the edges of the diagram, we have noted the additive error between the expressions at the nodes.

For reference, here are some useful definitions which we will use in the proof to follow.

$$p(w; \sigma) := \frac{1}{\sigma \sqrt{2\pi}} e^{-w^2/(2\sigma^2)} \quad (\text{C.1})$$

$$x(w) := \sqrt{p(w; \sigma)} \quad (\text{C.2})$$

$$X(f) := \mathcal{F}\{x\}(f) = \int_{\mathbb{R}} x(t) e^{-2\pi i t f} dt \quad (\text{C.3})$$

$$X_{1/T}\left(\frac{k}{2^q T}\right) := \sum_{n=-\infty}^{\infty} x(nT) \cdot e^{-i2\pi k n / 2^q} \quad k = -2^{q-1}, \dots, 2^{q-1} - 1 \quad (\text{C.4})$$

$$\mathcal{N}(\sigma, T, m) := \sum_{n=-(2^{m-1}-1)}^{2^{m-1}-1} |x(nT)|^2 \quad (\text{C.5})$$

$$x[n] := \begin{cases} \frac{1}{\mathcal{N}(\sigma, T, m)^{1/2}} x(nT) & n = -2^{m-1} - 1, \dots, 2^{m-1} - 1 \\ 0 & n = -2^{m-1} \end{cases} \quad (\text{C.6})$$

$$X[k] := \frac{1}{2^{q/2}} \sum_{n=-(2^{m-1}-1)}^{2^{m-1}-1} x[n] \cdot e^{-i2\pi k n / 2^q} \quad (\text{C.7})$$

It will also be convenient to define the Fourier dual sampling rates and widths, respectively.

$$F := (2^q T)^{-1} \quad (\text{C.8})$$

$$\sigma_f := (4\pi\sigma)^{-1} \quad (\text{C.9})$$

C.1 Truncation Error

The error between the scaled Discrete-time Fourier Transform (DTFT)

$$\frac{X_{1/T}\left(\frac{k}{2^q T}\right)}{\sqrt{2^q \mathcal{N}(\sigma, T, m)}} \quad (\text{C.10})$$

and the untruncated DFT $X[k]$ of Equation (C.7) can be estimated to be

$$\begin{aligned} \epsilon_{\text{trunc}} &= \frac{1}{2^{q/2} \mathcal{N}(\sigma, T, m)^{1/2}} \left(\sum_{n=2^{m-1}}^{\infty} \sqrt{p(nT; \sigma)} + \sum_{n=-\infty}^{-2^{m-1}} \sqrt{p(nT; \sigma)} \right) \\ &= \frac{2\tilde{\epsilon}_{\text{trunc}}}{2^{q/2} \mathcal{N}(\sigma, T, m)^{1/2}}, \end{aligned} \quad (\text{C.11})$$

where

$$\tilde{\epsilon}_{\text{trunc}} := \sum_{n=2^{m-1}}^{\infty} \sqrt{p(nT; \sigma)}. \quad (\text{C.12})$$

We can upper bound $\tilde{\epsilon}_{\text{trunc}}$ using a right-Riemann sum approximation of the error function,

$$\sum_{n=2^{m/2}}^{\infty} \sqrt{p(nT; \sigma)} \leq \sqrt[4]{\frac{\pi}{2}} \frac{\sqrt{\sigma}}{T} \operatorname{erfc} \left(\frac{T(2^m/2 - 1)}{2\sigma} \right)$$

and together with the simple bound [46]

$$\operatorname{erfc} x \leq e^{-x^2} \quad (x > 0) \quad (\text{C.13})$$

we have

$$\tilde{\epsilon}_{\text{trunc}} \leq \sqrt[4]{\frac{\pi}{2}} \frac{\sqrt{\sigma}}{T} \exp \left\{ - \left(\frac{T(2^{m-1} - 1)}{2\sigma} \right)^2 \right\}. \quad (\text{C.14})$$

Finally, we can express the truncation error ϵ_{trunc} , using (C.11) and the fact that $\mathcal{N}(\sigma, T, m) \in O(1/T)$.

$$\epsilon_{\text{trunc}} \leq \frac{2^{3/4} \pi^{1/4} \sqrt{\sigma}}{2^{q/2} \mathcal{N}(\sigma, T, m)^{1/2}} e^{-\left(\frac{T(2^{m-1}-1)}{2\sigma}\right)^2} \in O \left(\sqrt{\frac{\sigma}{2^q T}} e^{-\Omega(2^m/(\sigma/T))} \right) \quad (\text{C.15})$$

along

C.2 Aliasing error

Now that we have estimated the error on the DTFT by time-domain truncation, we would also like to estimate the aliasing error, that is, the difference between the DTFT and the Fourier transform over a period $1/T$. This is coming from having a finite sampling rate of a function that is not bounded in the frequency domain. In order to estimate the aliasing error, we must look at the Fourier transform of a Gaussian distribution, which is another Gaussian.

$$X(f) = \sqrt{p(f; \sigma_f)} \quad (\text{C.16})$$

Now, the DTFT can be expressed in terms of a Fourier transform in the following way

$$X_{1/T}(f) = \frac{1}{T} \sum_{k=-\infty}^{\infty} X(f - k/T). \quad (\text{C.17})$$

Therefore, we define the aliasing error as

$$\tilde{\epsilon}_{\text{alias}} := X_{1/T}(f) - \frac{1}{T}X(f) = \frac{1}{T} \sum_{k=-\infty}^{-1} X(f - k/T) + \frac{1}{T} \sum_{k=1}^{\infty} X(f - k/T). \quad (\text{C.18})$$

The aliasing error has a critical point at $f = 0$ and its second derivative is strictly positive throughout $f = [-1/2T, 1/2T]$ and thus we know that the error is largest at the boundaries of the DTFT. That is, at $f = -1/2T, 1/2T$. The error at these two points is expected to be the same so we just choose $f = 1/2T$ in order to bound the error from above. That means

$$\begin{aligned} \tilde{\epsilon}_{\text{alias}} &\leq \frac{1}{T} \sum_{k=-\infty}^{-1} X(1/2T - k/T) + \frac{1}{T} \sum_{k=1}^{\infty} X(1/2T - k/T) \\ &\leq \frac{1}{T} \sum_{k=-\infty}^{\infty} X(1/2T - k/T) \\ &= \pi^{1/4} 2^{3/4} \sqrt{\sigma} \frac{1}{T} \theta_2(e^{-(2\sigma\pi/T)^2}) \\ &= \pi^{1/4} 2^{3/4} \sqrt{\sigma} \frac{2}{T} e^{-(\pi\sigma/T)^2} \sum_{n=0}^{\infty} (e^{-4(2\pi\sigma/T)^2})^{n(n+1)}, \end{aligned} \quad (\text{C.19})$$

where $\theta_2(q) \equiv \theta_2(0, q)$ is the 2nd Jacobi theta function [47]. For $e^{-(2\pi\sigma/T)^2} \leq 1$,

$$\begin{aligned} \tilde{\epsilon}_{\text{alias}} &\leq \pi^{1/4} 2^{3/4} \sqrt{\sigma} \frac{1}{T} \cdot \left(2e^{-\pi^2(\sigma/T)^2} \sum_{n=0}^{\infty} (e^{-4\pi^2(\sigma/T)^2})^n \right) \\ &= \pi^{1/4} 2^{3/4} \sqrt{\sigma} \frac{1}{T} \cdot \left(2e^{-\pi^2(\sigma/T)^2} \frac{1}{1 - e^{-4\pi^2(\sigma/T)^2}} \right) \end{aligned} \quad (\text{C.20})$$

For $e^{-4\pi^2(\sigma/T)^2} \leq 1/2$

$$\begin{aligned}\tilde{\epsilon}_{\text{alias}} &\leq \pi^{1/4} 2^{3/4} \sqrt{\sigma} \frac{1}{T} \cdot \left(2e^{-\pi^2(\sigma/T)^2} \frac{1}{1 - e^{-4\pi^2(\sigma/T)^2}} \right) \\ &\leq 4\pi^{1/4} 2^{3/4} \sqrt{\sigma} \frac{1}{T} \cdot e^{-\pi^2(\sigma/T)^2}\end{aligned}\tag{C.21}$$

Therefore, the error ϵ_{alias} between

$$\frac{X_{1/T}(k/(2^q T))}{2^{q/2} \mathcal{N}(\sigma, T, m)^{1/2}}\tag{C.22}$$

and

$$\frac{X(k/(2^q T))}{T 2^{q/2} \mathcal{N}(\sigma, T, m)^{1/2}}\tag{C.23}$$

can be estimated using (C.21) to be

$$\epsilon_{\text{alias}} = \frac{\tilde{\epsilon}_{\text{alias}}}{2^{q/2} \mathcal{N}(\sigma, T, m)^{1/2}} \leq \left(\frac{4\pi^{1/4} 2^{3/4} \sqrt{\sigma} \frac{1}{T} \cdot e^{-\pi^2(\sigma/T)^2}}{2^{q/2} \mathcal{N}(\sigma, T, m)^{1/2}} \right).\tag{C.24}$$

Thus, with a choice of

$$\frac{\sigma}{T} \sim \sqrt{2^m},\tag{C.25}$$

both sources of error can be bounded with

$$\begin{aligned}\epsilon_{\text{trunc}} &= O\left(\frac{2^{m/4}}{2^{q/2} e^{\Omega(2^m)}}\right) \\ \epsilon_{\text{alias}} &= O\left(\frac{2^{m/4}}{2^{q/2} e^{\Omega(2^m)}}\right).\end{aligned}\tag{C.26}$$

This means that after the DFT, we know that we will have in the register the amplitude

$$\sqrt{\frac{F}{T \mathcal{N}(\sigma, T, m)}} X(kF) + \epsilon_{\text{DFT}},\tag{C.27}$$

where ϵ_{DFT} is the total error from both truncation as well as aliasing. We then bound the error by summing (C.21) and (C.10) to find

$$\begin{aligned}\epsilon_{\text{DFT}} &\leq \frac{(8\pi)^{1/4} \sqrt{\sigma}}{T \sqrt{2^q \mathcal{N}(\sigma, T, m)}} e^{-\left(\frac{T(2^{m-1}-1)}{2\sigma}\right)^2} + 4 \frac{(8\pi)^{1/4} \sqrt{\sigma}}{T \sqrt{2^q \mathcal{N}(\sigma, T, m)}} e^{-\pi^2(\sigma/T)^2} \\ &\leq 2 \max(\epsilon_{\text{alias}}, \epsilon_{\text{trunc}}) \in O\left(\frac{2^{m/4}}{2^{q/2} e^{\Omega(2^m)}}\right).\end{aligned}\tag{C.28}$$

C.3 Renormalization error and ultimate additive error

We know that the signal after DFT has to be normalized, we also know that it is locally ϵ_{DFT} close to Equation (C.23), which is not necessarily normalized. Therefore, for the norm of samples of Equation (C.23) we obtain the following upper bound on normalization ratios

$$\begin{aligned}
\frac{F}{T\mathcal{N}(\sigma, T, m)} \sum_k |X(kF)|^2 &= \frac{F\mathcal{N}(\sigma_f, F, q)}{T\mathcal{N}(\sigma, T, m)} \\
&= 1 - 2 \left(\frac{F}{T\mathcal{N}(\sigma, T, m)} \right)^{1/2} \Re(\vec{\epsilon}_{\text{DFT}} \cdot \vec{X}) + \|\vec{\epsilon}_{\text{DFT}}\|^2 \\
&\leq 1 + \|\vec{\epsilon}_{\text{DFT}}\|^2 \\
&\leq 1 + \|\vec{\epsilon}_{\text{DFT}}\| \\
&\leq 1 + \epsilon_{\text{DFT}} 2^{q/2},
\end{aligned} \tag{C.29}$$

where we have used the assumption $\|\vec{\epsilon}_{\text{DFT}}\| \leq 1$. Similarly, the lower bound is

$$\begin{aligned}
\frac{F}{T\mathcal{N}(\sigma, T, m)} \sum_k \left| X\left(\frac{k}{2^q T}\right) \right|^2 &= \frac{F\mathcal{N}(\sigma_f, F, q)}{T\mathcal{N}(\sigma, T, m)} \\
&\geq 1 - 2 \left(\frac{F}{T\mathcal{N}(\sigma, T, m)} \right)^{1/2} \Re(\|\vec{\epsilon}_{\text{DFT}}\| \|\vec{X}\|) \\
&\geq 1 - 2\epsilon_{\text{DFT}} 2^{q/2} \left(\frac{F\mathcal{N}(\sigma_f, F, q)}{T\mathcal{N}(\sigma, T, m)} \right)^{1/2} \\
&\geq 1 - 2\epsilon_{\text{DFT}} 2^{q/2} (1 + \|\vec{\epsilon}_{\text{DFT}}\|^2) \\
&\geq 1 - 4\epsilon_{\text{DFT}} 2^{q/2}.
\end{aligned} \tag{C.30}$$

Thus, if we take the full inequality

$$1 - 4\epsilon_{\text{DFT}} 2^{q/2} \leq \left(\frac{F\mathcal{N}(\sigma_f, F, q)}{T\mathcal{N}(\sigma, T, m)} \right) \leq 1 + \epsilon_{\text{DFT}} 2^{q/2} \tag{C.31}$$

and take the square root, we obtain,

$$1 - 2\epsilon_{\text{DFT}} 2^{q/2} \leq \left(\frac{F\mathcal{N}(\sigma_f, F, q)}{T\mathcal{N}(\sigma, T, m)} \right)^{1/2} \leq 1 + \epsilon_{\text{DFT}} 2^{q/2}, \tag{C.32}$$

where we have used $1 + \sqrt{x} \leq 1 + x$ and $1 - \sqrt{x} \geq 1 - x/2$ for $x \leq 1$. We now divide every side by $\mathcal{N}(\sigma_f, F, q)$ to obtain

$$1/\mathcal{N}(\sigma_f, F, q)^{1/2} - 2\epsilon_{\text{DFT}} 2^{q/2}/\mathcal{N}(\sigma_f, F, q) \leq \left(\frac{F}{T\mathcal{N}(\sigma, T, m)} \right)^{1/2} \leq$$

$$1/\mathcal{N}(\sigma_f, F, q)^{1/2} + \epsilon_{\text{DFT}}\sqrt{2^q}/\mathcal{N}(\sigma_f, F, q)^{1/2}. \quad (\text{C.33})$$

Subtract $1/\mathcal{N}(\sigma_f, F, q)$, which gives

$$\begin{aligned} -2\epsilon_{\text{DFT}}\sqrt{\frac{2^q}{\mathcal{N}(\sigma_f, F, q)}} &\leq \sqrt{\frac{F}{T\mathcal{N}(\sigma, T, m)}} - \sqrt{\frac{1}{\mathcal{N}(\sigma_f, F, q)}} \\ &\leq \epsilon_{\text{DFT}}\sqrt{\frac{2^q}{\mathcal{N}(\sigma_f, F, q)}}. \end{aligned} \quad (\text{C.34})$$

We multiply these inequalities by $\max |X(k)|$

$$\begin{aligned} -2\epsilon_{\text{DFT}}\sqrt{\frac{2^q}{\mathcal{N}(\sigma_f, F, q)}} \left(\frac{1}{2\pi\sigma_f^2}\right)^{1/4} \\ \leq \left(\sqrt{\frac{F}{T\mathcal{N}(\sigma, T, m)}} - \sqrt{\frac{1}{\mathcal{N}(\sigma_f, F, q)}}\right) \max |X(k)| \\ \leq \epsilon_{\text{DFT}}\sqrt{\frac{2^q}{\mathcal{N}(\sigma_f, F, q)}} \left(\frac{1}{2\pi\sigma_f^2}\right)^{1/4}. \end{aligned} \quad (\text{C.35})$$

We take the absolute value to have a two-sided inequality, and use the definitions (C.8) and (C.9).

$$\begin{aligned} &\left|\sqrt{\frac{F}{T\mathcal{N}(\sigma, T, m)}} - \sqrt{\frac{1}{\mathcal{N}(\sigma_f, F, q)}}\right| \max |X(k)| \\ &\leq 2\epsilon_{\text{DFT}}\sqrt{\frac{1}{TF\mathcal{N}(\sigma_f, F, q)}} \left(\frac{(4\pi\sigma)^2}{2\pi}\right)^{1/4} \\ &\leq 2\frac{(8\pi)^{1/4}\sqrt{\sigma}}{T\sqrt{2^q\mathcal{N}(\sigma, T, m)}}\sqrt{\frac{1}{TF\mathcal{N}(\sigma_f, F, q)}} \left(\frac{(4\pi\sigma)^2}{2\pi}\right)^{1/4} \\ &\quad \max \left(4 \exp \left\{ -\pi^2(\sigma/T)^2 \right\}, \exp \left\{ -\left(\frac{2^{m-1}-1}{2\sigma/T}\right)^2 \right\} \right) \\ &\leq 4\frac{\sigma}{T}\sqrt{\frac{2\pi}{2^qT\mathcal{N}(\sigma, T, m)F\mathcal{N}(\sigma_f, F, q)}} \\ &\quad \max \left(4 \exp \left\{ -\pi^2(\sigma/T)^2 \right\}, \exp \left\{ -\left(\frac{2^{m-1}-1}{2\sigma/T}\right)^2 \right\} \right) \end{aligned} \quad (\text{C.36})$$

It is worth noting that, for the purpose of estimating $(T\mathcal{N}(\sigma, T, m))^{1/2}$, we can increase 2^q arbitrarily (or decrease F arbitrarily). Through this, we see that:

$$\lim_{F \rightarrow 0} F\mathcal{N}(\sigma_f, F, q) = \lim_{F \rightarrow 0} \text{erf} \left(\frac{(2^{q-1} - 1/2)F}{\sqrt{2}\sigma_f} \right) = \text{erf} \left(\sqrt{2}\pi\sigma/T \right)$$

$$\geq 1 - \frac{1}{\exp\left(\sqrt{2\pi}\sigma/T\right)^2} \quad (\text{C.37})$$

For our choice of $\frac{\sigma}{T}$ this would mean that

$$(T\mathcal{N}(\sigma, T, m))^{1/2} = 1 - O\left(\frac{2^{m/4}}{\exp(\Omega(2^m))}\right), \quad (\text{C.38})$$

and

$$\begin{aligned} T^{1/2}\mathcal{N}(\sigma, T, m)^{1/2} &= F^{1/2}\mathcal{N}(\sigma_f, F, q)^{1/2} \\ &\quad - T^{1/2}\mathcal{N}(\sigma, T, m)^{1/2}O\left(\frac{2^{m/4}}{\exp(\Omega(2^m))}\right) \\ &= F^{1/2}\mathcal{N}(\sigma_f, F, q)^{1/2} - O\left(\frac{2^{m/4}}{\exp(\Omega(2^m))}\right). \end{aligned} \quad (\text{C.39})$$

Therefore, the amplitudes that we get on the ancillary register are

$$\frac{1}{\sqrt{\mathcal{N}(\sigma_f, F, q)}}X\left(\frac{k}{2^q T}\right) + O\left(\frac{2^{(m-q)/2}}{\exp(\Omega(2^m))}\right) + \epsilon_{\text{DFT}}. \quad (\text{C.40})$$

D Proofs of State Preparation and Ground State Energy Extrapolation Costs

Here we prove results related to the costs of state preparation and phase estimation for extrapolation of the ground state energy, from Theorem 15 and Theorem 17 respectively.

D.1 Proof of Lemma 12

Proof. Let λ_{m-1} , λ_m , and λ_{m+1} be consecutive simple eigenvalues of H

$$\lambda_{m-1} < \lambda_m < \lambda_{m+1} \quad (\text{D.1})$$

such that $|\lambda_m - \lambda_{m-1}| \geq \gamma_0$, and $|\lambda_m - \lambda_{m+1}| \geq \gamma_0$. With $\lambda(\tau)$ as in the lemma statement, we have $\lambda_m(0) = \lambda_m$ for all indices m . By Lemma 11, there exist eigenvalues $\lambda_j(\tau)$, $\lambda_k(\tau)$, and $\lambda_l(\tau)$ of \tilde{H}_z such that

$$\begin{aligned} \max_{\tau \in \{\tau: |\tau| \leq r\}} |\lambda_{m-1}(0) - \lambda_j(\tau)| &\leq \max_{\tau \in \{\tau: |\tau| \leq r\}} \|H - \tilde{H}_z\| \\ \max_{\tau \in \{\tau: |\tau| \leq r\}} |\lambda_m(0) - \lambda_k(\tau)| &\leq \max_{\tau \in \{\tau: |\tau| \leq r\}} \|H - \tilde{H}_z\| \\ \max_{\tau \in \{\tau: |\tau| \leq r\}} |\lambda_{m+1}(0) - \lambda_l(\tau)| &\leq \max_{\tau \in \{\tau: |\tau| \leq r\}} \|H - \tilde{H}_z\| \end{aligned} \quad (\text{D.2})$$

for some j , k , and l . We are guaranteed that $j \neq m$, $l \neq m$ and $k = m$ if

$$\max_{\tau \in \{\tau: |\tau| \leq r\}} \|H - \tilde{H}_z\| \leq \gamma_0/2. \quad (\text{D.3})$$

Assuming we have the upper bound

$$\|H - \tilde{H}_z\| \leq \alpha \frac{|\tau|^p}{(p+1)!} e^{\beta|\tau|}, \quad (\text{D.4})$$

this guarantee can be satisfied by the constraint

$$\gamma_0/2 \geq \max_{\tau \in \{\tau: |\tau| \leq r\}} \left(\alpha \frac{|\tau|^p}{(p+1)!} e^{\beta|\tau|} \right) = \alpha \frac{r^p}{(p+1)!} e^{\beta r}. \quad (\text{D.5})$$

Solving for r on the last inequality gives us

$$r \leq \frac{p}{\beta} \text{LambertW} \left(\frac{\beta}{p} \left(\frac{\gamma_0(p+1)!}{2\alpha} \right)^{1/p} \right). \quad (\text{D.6})$$

If λ_m is the ground state or highest state of H , we just repeat a similar proof except there is only one state above or below respectively. \square

D.2 Proof of Lemma 13

Proof. We take a look at the complex-time Schrödinger equation for $S_p(\tau)$.

$$S'_p(\tau) = \overbrace{i(H + \mathcal{E}(\tau))}^A S_p(\tau) \quad (\text{D.7})$$

We now recall that we can express

$$\Omega(\tau) = \log S_p(\tau) \quad (\text{D.8})$$

through the Magnus expansion for the complex-time derivative,

$$\Omega'(\tau) = \sum_{n=0}^{\infty} \frac{B_n}{n!} \text{ad}_{\Omega(\tau)}^n(A). \quad (\text{D.9})$$

We then re-express $E(\tau)$ and expand Ω'

$$\begin{aligned} E(\tau) &= \int_{\mathcal{P}} \Omega'(\tau_1) d\tau_1 - \int_{\mathcal{P}} H d\tau_1 \\ &= \int_{\mathcal{P}} (\Omega'(\tau_1) - H) d\tau_1 \\ &= \int_{\mathcal{P}} \left(\mathcal{E}(\tau_1) + \sum_{n=1}^{\infty} \frac{B_n}{n!} \text{ad}_{\Omega(\tau_1)}^n(A(\tau_1)) \right) d\tau_1. \end{aligned} \quad (\text{D.10})$$

Here, \mathcal{P} is any path in the complex plane going from $\tau_1 = 0$ to $\tau_1 = \tau$. We now expand out the first application of ad_Ω on A and separate the $n = 1$ term from the sum obtaining:

$$E(\tau) = \int_{\mathcal{P}} \left(\mathcal{E}(\tau_1) + B_1[\Omega(\tau_1), A(\tau_1)] + \sum_{n=2}^{\infty} \frac{B_n}{n!} \text{ad}_\Omega^{n-1}[\Omega(\tau_1), A(\tau_1)] \right) d\tau_1. \quad (\text{D.11})$$

Following this, we make use of the estimation lemma ('ML' inequality) to obtain

$$\|E(\tau)\| \leq \int_0^S |\gamma_{\mathcal{P}}| \left(\|\mathcal{E}\| + |B_1| \|[\Omega, A]\| + \sum_{n=2}^{\infty} \frac{|B_n|}{n!} \|\text{ad}_\Omega^{n-1}\| \|[\Omega, A]\| \right) ds \quad (\text{D.12})$$

where s parametrizes the path \mathcal{P} and $|\gamma_{\mathcal{P}}| ds = \left| \frac{d\tau_1}{ds} \right| ds$ is the arc-length differential. Now, we make use of the bound $\|\text{ad}_\Omega\| \leq 2\|\Omega\|$ to obtain:

$$\begin{aligned} &\leq \int_0^S |\gamma_{\mathcal{P}}| \left(\|\mathcal{E}\| + |B_1| \|[\Omega, A]\| + \sum_{n=2}^{\infty} \frac{|B_n|}{n!} (2\|\Omega\|)^{n-1} \|[\Omega, A]\| \right) ds \\ &\leq \int_0^S |\gamma_{\mathcal{P}}| \left(\|\mathcal{E}\| + |B_1| \|[\Omega, A]\| + \frac{\|[\Omega, A]\|}{(2\|\Omega\|)} \sum_{n=2}^{\infty} \frac{|B_n|}{n!} (2\|\Omega\|)^n \right) ds \end{aligned} \quad (\text{D.13})$$

We now make use of

$$\sum_{n=2}^{\infty} |B_n| \frac{|t|^n}{n!} = \sum_{n=1}^{\infty} B_n^* \frac{|t|^{2n}}{(2n)!} = 1 - \frac{|t|}{2} \cot \frac{|t|}{2} \quad \text{for } |t| \leq 2\pi \quad (\text{D.14})$$

to obtain:

$$\begin{aligned} \|E(\tau)\| &\leq \int_0^S |\gamma_{\mathcal{P}}| \left(\|\mathcal{E}\| + \frac{1}{2} \|[\Omega, A]\| + \frac{\|[\Omega, A]\|}{(2\|\Omega\|)} (1 - \|\Omega\| \cot \|\Omega\|) \right) ds \\ &\leq \int_0^S |\gamma_{\mathcal{P}}| \left(\|\mathcal{E}\| + \frac{1}{2} \|[\Omega, A]\| + \frac{1}{2} \|[\Omega, A]\| \left(\frac{1}{\|\Omega\|} - \cot \|\Omega\| \right) \right) ds \end{aligned} \quad (\text{D.15})$$

We now assume that

$$|\gamma_{\mathcal{P}}| = \left| \frac{d\tau}{ds} \right| = 1, \quad (\text{D.16})$$

and

$$s = |\tau_1|, \quad (\text{D.17})$$

thus

$$\|E(\tau)\| \leq \int_0^{|\tau|} \left(\|\mathcal{E}\| + \frac{1}{2} \|\Omega, A\| + \frac{1}{2} \|\Omega, A\| \left(\frac{1}{\|\Omega\|} - \cot \|\Omega\| \right) \right) ds. \quad (\text{D.18})$$

We assume

$$\left(\frac{1}{\|\Omega\|} - \cot \|\Omega\| \right) \leq 1, \quad (\text{D.19})$$

which is fulfilled provided

$$\|E(\tau)\| \leq 1. \quad (\text{D.20})$$

Now,

$$\|E(\tau)\| \leq \int_0^{|\tau|} (\|\mathcal{E}\| + \|\Omega, A\|) ds. \quad (\text{D.21})$$

Recalling that $\Omega(\tau) = E(\tau) + \tau H$,

$$\|\Omega, A\| = \|[E + \tau H, \mathcal{E} + H]\| \leq 2\|E\| (\|\mathcal{E}\| + \|H\|) + 2|\tau| \|H\| \|\mathcal{E}\|. \quad (\text{D.22})$$

Plugging back into Equation (D.21) and using estimation lemma for $\|E\|$ inside the integral we obtain:

$$\|E(\tau)\| \leq \int_0^{|\tau|} (\|\mathcal{E}\| + 2s\|H\|\|\mathcal{E}\|) ds + 2 \max_{|\tilde{\tau}| \leq |\tau|} \|E(\tilde{\tau})\| \int_0^{|\tau|} (\|\mathcal{E}\| + \|H\|) ds. \quad (\text{D.23})$$

We now do the relabelling $\tau \rightarrow \tau_2$ and then take $\max_{|\tau_2| \leq |\tau|}$ of both sides

$$\max_{|\tau_2| \leq |\tau|} \|E(\tau_2)\| \leq \max_{|\tau_2| \leq |\tau|} \int_0^{|\tau_2|} (\|\mathcal{E}\| + 2s\|H\|\|\mathcal{E}\|) ds + 2 \max_{|\tau_2| \leq |\tau|} \|E(\tau_2)\| \max_{|\tau_2| \leq |\tau|} \int_0^{|\tau_2|} (\|\mathcal{E}\| + \|H\|) ds, \quad (\text{D.24})$$

which becomes

$$\max_{|\tau_2| \leq |\tau|} \|E(\tau_2)\| \leq \int_0^{|\tau|} (\|\mathcal{E}\| + 2s\|H\|\|\mathcal{E}\|) ds + 2 \max_{|\tau_2| \leq |\tau|} \|E(\tau_2)\| \int_0^{|\tau|} (\|\mathcal{E}\| + \|H\|) ds, \quad (\text{D.25})$$

Solving for $\max_{|\tau_2| \leq |\tau|} \|E(\tau_2)\|$

$$\begin{aligned} \max_{|\tau_2| \leq |\tau|} \|E(\tau_2)\| &\leq \frac{\int_0^{|\tau|} (\|\mathcal{E}\| + 2s\|H\|\|\mathcal{E}\|) ds}{1 - 2|\tau| \|H\| - 2 \int_0^{|\tau|} \|\mathcal{E}\| ds} \\ &\leq \frac{(1 + 2|\tau| \|H\|) \int_0^{|\tau|} (\|\mathcal{E}\|) ds}{1 - 2|\tau| \|H\| - 2 \int_0^{|\tau|} \|\mathcal{E}\| ds} \end{aligned}$$

(D.26)

We now assume

$$\int_0^{|\tau|} \|\mathcal{E}\| ds \leq 1/8 \quad (\text{D.27})$$

and

$$|\tau| \leq 1/8, \quad (\text{D.28})$$

thus,

$$\|E(\tau)\| \leq \max_{|\tau_2| \leq |\tau|} \|E(\tau_2)\| \leq \frac{5}{2} \int_0^{|\tau|} \|\mathcal{E}(\tau_1)\| ds. \quad (\text{D.29})$$

□

D.3 Proof of Lemma 14

Our task here is to show that a bound of the form shown (3.17) actually exists, and for this we follow closely the work done in [18]. In order to make assertions over the general complex plane we will extend their results to complex time. The extension will be straightforward because

$$e^{i\tau A}$$

has no singularities except for non-removable singularities at $\text{Im}(\tau) \in \{-\infty, \infty\}$ for a Hermitian A and at $\text{Im}(\tau) = -\infty$ when it is positive semi-definite. Thus, we can replace all the real-axis integrals

$$\int_0^t f(t_1) dt_1$$

for the contour integrals

$$\int_C f(\tau_1) d\tau_1,$$

where \mathcal{P} is any path from 0 to a generally complex τ as long as that path does not include $\text{Im}(\tau) \in \{-\infty, \infty\}$. Our results are summarized by the following

Proof. First, we define a general product formula the following way

$$\mathcal{S}(\tau) := \prod_{v=1}^{\Upsilon} \prod_{m=1}^m e^{i\tau a_{(v,m)} H_{\pi_v(m)}}, \quad (\text{D.30})$$

where the coefficients $a_{(v,m)}$ are real numbers. The parameter Υ denotes the number of *stages* of the product formula. For example, for the ST formula $S_{2k}(t)$, we have $\Upsilon = 2 \cdot 5^{k-1}$. The permutation π_v controls the ordering of operator summands within

stage v of the product formula. We then consider the error in the effective Hamiltonian with the help of the definition of the path-ordered exponential:

$$\mathcal{S}(\tau) = \exp_{\mathcal{P}} \left(i \int d\tau (H + \mathcal{E}(\tau)) \right). \quad (\text{D.31})$$

Using this definition and Lemma 13, we can write the operator error on the Hamiltonian the following way:

$$\left\| \tilde{H}_z(t) - H \right\| \leq \frac{5}{2} \left(\int_0^{|\tau|} \|\mathcal{E}(\tau_1)\| ds \right) / |\tau|. \quad (\text{D.32})$$

where s in this context is $|\tau_1|$. Let \succ be a lexicographic, or dictionary, ordering on tuples (v, m) , i.e. $(v, m) \succ (v', m')$ if $v > v'$, or if $v = v'$ and $m \geq m'$. Let

$$\mathcal{S}(\tau) = \overleftarrow{\prod}_{(v,m)} e^{i\tau a_{(v,m)} H_{\pi_v(m)}}, \quad (\text{D.33})$$

be a product formula with $\overleftarrow{\prod}_{(v,m)}$ dictating that the operations are applied from right to left in ascending order according to \succ . Let $\overrightarrow{\prod}_{(v,m)}$ refer to descending order. After differentiating the evolution operator, and some algebra, one obtains that $\mathcal{E}(\tau)$ can be expressed as

$$\mathcal{E}(\tau) = \sum_{(v,m)} \overleftarrow{\prod}_{(v',m') \succ (v,m)} e^{i\tau a_{(v',m')} H_{\pi_{v'}(m')}} \left(a_{(v,m)} H_{\pi_v(m)} \right) \overrightarrow{\prod}_{(v',m') \succ (v,m)} e^{-i\tau a_{(v',m')} H_{\pi_{v'}(m')}} - H. \quad (\text{D.34})$$

We can bound the norm of this term through the following Taylor expansion of the term

$$\begin{aligned} & e^{i\tau A_s} \dots e^{i\tau A_2} e^{i\tau A_1} B e^{-i\tau A_1} e^{-i\tau A_2} \dots e^{-i\tau A_s} \\ &= C_0 + C_1 \tau + \dots + C_{p-1} \tau^{p-1} \\ &+ i^p \sum_{k=1}^s \sum_{\substack{q_1 + \dots + q_k = p \\ q_k \neq 0}} e^{i\tau A_s} \dots e^{i\tau A_{k+1}} \\ &\quad \cdot \int_C d\tau_2 e^{i\tau_2 A_k} \text{ad}_{A_k}^{q_k} \dots \text{ad}_{A_1}^{q_1} (B) e^{-i\tau_2 A_k} \cdot \frac{(\tau - \tau_2)^{q_k-1} \tau^{q_1 + \dots + q_{k-1}}}{(q_k - 1)! q_{k-1}! \dots q_1!} \\ &\quad \cdot e^{-i\tau A_{k+1}} \dots e^{-i\tau A_s}, \end{aligned} \quad (\text{D.35})$$

where $\text{ad}_A(B) := [A, B]$. We can bound the norm of that last term in Equation (D.35) through triangle inequality for contour integrals, and for sake of simplicity, we fix contour C going straight from 0 to τ . We also assume for the moment that matrices A_j are positive definite. Thus, the upper bound becomes:

$$\sum_{k=1}^s \sum_{\substack{q_1 + \dots + q_k = p \\ q_k \neq 0}} \prod_{j=k+1}^s \overbrace{\|e^{i\tau A_j}\| \|e^{-i\tau A_j}\|}^{\leq e^{2|\text{Im}(\tau)| \|A_j\|}} \frac{|\tau|^{q_1 + \dots + q_{k-1}}}{(q_k - 1)! q_{k-1}! \dots q_1!} \left\| \text{ad}_{A_k}^{q_k} \dots \text{ad}_{A_1}^{q_1} (B) \right\|$$

$$\int_C d\tau_2 |\tau - \tau_2|^{q_k-1} \overbrace{\|e^{i\tau_2 A_k}\| \|e^{-i\tau_2 A_k}\|}^{\leq e^{2|\text{Im}(\tau_2)| \|A_k\|}} \quad (\text{D.36})$$

$$\leq \sum_{k=1}^s \sum_{\substack{q_1+\dots+q_k=p \\ q_k \neq 0}} \binom{p}{q_1 \dots q_k} \frac{|\tau|^p}{p!} \|\text{ad}_{A_k}^{q_k} \dots \text{ad}_{A_1}^{q_1}(B)\| e^{2|\tau| \sum_{j=k}^s \|A_j\|} \quad (\text{D.37})$$

$$\leq \sum_{q_1+\dots+q_s=p} \binom{p}{q_1 \dots q_s} \frac{|\tau|^p}{p!} \|\text{ad}_{A_s}^{q_s} \dots \text{ad}_{A_1}^{q_1}(B)\| e^{2|\tau| \sum_{j=1}^s \|A_j\|} \quad (\text{D.38})$$

$$= \alpha_{\text{comm}}(A_s, \dots, A_1, B) \frac{|\tau|^p}{p!} e^{2|\tau| \sum_{j=1}^s \|A_j\|}, \quad (\text{D.39})$$

where

$$\alpha_{\text{comm}}(A_s, \dots, A_1, B) := \sum_{q_1+\dots+q_s=p} \binom{p}{q_1 \dots q_s} \|\text{ad}_{A_s}^{q_s} \dots \text{ad}_{A_1}^{q_1}(B)\|. \quad (\text{D.40})$$

Thus, the upper bound on $\|\mathcal{E}(\tau)\|$ is

$$\begin{aligned} \|\mathcal{E}(\tau)\| &\leq \sum_{(v,m)} \alpha_{\text{comm}}(a_{(\Upsilon,m)} H_{\pi_{\Upsilon}(m)}, \dots, a_{(v,m+1)} H_{\pi_v(m+1)}, a_{(v,m)} H_{\pi_v(m)}) \\ &\quad \cdot \frac{|\tau|^p}{p!} e^{2|\tau| \sum_{(v',m') \succ (v,m)} \|a_{(v',m')} H_{\pi_{v'}(m')}\|}, \end{aligned} \quad (\text{D.41})$$

where $(v, m+1) = (v+1, 1)$. With this, the upper bound in $\|\tilde{H}_z(t) - H\|$

$$\begin{aligned} \|H - \tilde{H}_z\| &\leq \frac{5}{2} \sum_{(v,m)} \alpha_{\text{comm}}(a_{(\Upsilon,m)} H_{\pi_{\Upsilon}(m)}, \dots, a_{(v,m+1)} H_{\pi_v(m+1)}, a_{(v,m)} H_{\pi_v(m)}) \\ &\quad \cdot \frac{|\tau|^p}{(p+1)!} e^{2|\tau| \sum_{(v',m') \succ (v,m)} \|a_{(v',m')} H_{\pi_{v'}(m')}\|} \\ &\leq \frac{5}{2} \sum_{(v,m)} \alpha_{\text{comm}}(a_{(\Upsilon,m)} H_{\pi_{\Upsilon}(m)}, \dots, a_{(v,m+1)} H_{\pi_v(m+1)}, a_{(v,m)} H_{\pi_v(m)}) \\ &\quad \cdot \frac{|\tau|^p}{(p+1)!} e^{2|\tau| \sum_j \|H_j\|} \end{aligned} \quad (\text{D.42})$$

□

D.4 Avoiding fractional queries

For the Gaussian QPEA defined in Section 3.2.3, we must define the \tilde{U}_{s_k} operator. A first attempt looks the following way

$$\tilde{U}_{s_k} = (S_p(t s_k))^{e_k} = \left(\exp \left\{ -i t s_k \tilde{H}_{s_k} \right\} \right)^{e_k}, \quad (\text{D.43})$$

where s_k are the Chebyshev nodes with $a = 1$, and

$$e_k = \frac{s_1}{s_k}, \quad (\text{D.44})$$

such that the total simulation times

$$T = s_1 t, \quad (\text{D.45})$$

which is the same for each \tilde{U}_{s_k} .

This, however, will require fractional powers of (that is, a non-integer number of queries of) $S_p(t s_k)$. Although this can be efficiently achieved through quantum signal processing and one extra ancillary qubit, within this section we would like to stay as rudimentary as possible, minimizing quantum overhead. Hence, we will instead use the evolution operator

$$\tilde{U}'_{s_k} = (S_p(t s_k))^{e'_k} = \exp \left\{ -i t s_k \operatorname{sgn} s_k \left[\frac{s_1}{|s_k|} \right] \tilde{H}_{s_k} \right\}. \quad (\text{D.46})$$

Here,

$$e'_k = \operatorname{sgn} s_k \left[\frac{s_1}{|s_k|} \right], \quad (\text{D.47})$$

is an integer by construction. For this new operator \tilde{U}'_{s_k} , the Trotter step size remains

$$dt = t s_k, \quad (\text{D.48})$$

but the total time evolution changes to

$$T'_k = e'_k \times dt = t s_k \operatorname{sgn} s_k \left[\frac{s_1}{|s_k|} \right] = \Theta(t s_1). \quad (\text{D.49})$$

We note that now the equivalent evolution time T'_k varies for different k 's, but this has a negligible effect in state preparation (See Theorem 15) or in eigen energy estimation (See Theorem 17) with respect to the effective Hamiltonian \tilde{H}_{s_k} .

When it comes to query cost, this has the effect of increasing the resolution and the cost of QPEA by a factor of

$$\frac{|e'_k|}{|e_k|} \leq 2. \quad (\text{D.50})$$

As a last step before jumping into the proof of Theorem 15 we will first discuss some estimates for the coefficients of the terms in the extrapolation. The first such result is given in In the framework we propose here, most algorithms' cost will scale proportionally to $\|e'\|_1 \leq 2\|e\|_1$.

D.5 Proof of Theorem 15

Proof of Theorem 15. In the framework we propose here, most algorithms' cost will scale from Lemma 3

$$\|e'\|_1 \leq 2\|e\|_1 = 2s_1 \sum_{k=1}^n \frac{1}{|s_k|} = O(n \log n). \quad (\text{D.51})$$

For example, the ground state preparation methods detailed in [36] would scale proportionally to $\|e\|_1$. More precisely, the cost of state preparation scales with the minimum eigenvalue gap as

$$C_{\text{prep}} = O\left(\frac{\|e'\|_1}{t \min_k (E_1(s_k) - E_0(s_k))}\right), \quad (\text{D.52})$$

where, from Lemma 11 and the fact that the spectral gap is bounded below by γ_0 , we have that

$$\begin{aligned} \min_k (E_1(s_k) - E_0(s_k)) &\geq \gamma_0 - \max_{t_1 \leq t, s \leq 1} \|H - \tilde{H}_s(t_1)\| \\ &\geq \gamma_0 - O(t^p). \end{aligned} \quad (\text{D.53})$$

Thus, for $t = O(\gamma_0^{1/p})$,

$$\min_k (E_1(s_k) - E_0(s_k)) = \Omega(\gamma_0). \quad (\text{D.54})$$

From Lemma 12, we also have that

$$r = \Omega\left(\left(\frac{(p+1)!}{2^{p+1}}\right)^{1/p} \frac{\gamma_0^{1/p}}{t}\right) \quad (\text{D.55})$$

Recalling Lemma 10 that the interpolation error

$$\epsilon = O\left(\frac{1}{\rho^{M+1}}\right) \quad (\text{D.56})$$

Thus, solving for n we find:

$$n = O\left(\frac{\log 1/\epsilon}{\log \rho}\right) \quad (\text{D.57})$$

We also have that the radius of the Bernstein ellipse needed satisfies $\rho_{\max} = \Theta(r_{\max})$ from Lemma 12; hence,

$$n = O\left(\frac{p \log(1/\epsilon)}{\log(\gamma_0/t^p) + \log((p+1)!/2^{p+1})}\right). \quad (\text{D.58})$$

It is clear that a constraint is

$$t = \Omega\left(\gamma_0^{1/p}\right), \quad (\text{D.59})$$

since we need the radius to be a quantity greater than 1 and there is no advantage in going to a much smaller t . Finally, the cost for state preparation is

$$C_{\text{prep}} = O\left(\frac{\log(1/\epsilon) \log \log(1/\epsilon)}{\gamma_0^{1+1/p}}\right). \quad (\text{D.60})$$

□

D.6 Proof of Theorem 17

Proof of Theorem 17. Now, the cost of estimating the energy through GPE would, naively speaking, also scale like $O(n \log n)$. However, to get a tighter cost bound, we will allow the variances of observables to vary across nodes and then minimize the cost function:

$$L_0 = \sum_{k=1}^{n/2} \frac{1}{\sigma_k \cos \frac{2k-1}{2n} \pi} \quad (\text{D.61})$$

with the constraint

$$2 \sum_{k=1}^{n/2} c_k^2 \sigma_k^2 - \sigma_P^2 = 0. \quad (\text{D.62})$$

Thus, we the cost function with Lagrange multiplier is:

$$L = L_0 + \lambda \left(2 \sum_{k=1}^{n/2} c_k^2 \sigma_k^2 - \sigma_P^2 \right). \quad (\text{D.63})$$

Note that the terms in the cost function L_0 go like $1/\sigma_k$ due to Fourier duality and Theorem 16 and the other factor is from the cost of a single U_k implementation. After doing the minimization we are left with:

$$\begin{aligned} L_0 &= \frac{2^{1/2}}{\sigma_P} \left(\sum_{k=1}^{n/2} \frac{|c_k|^{2/3}}{\cos^{2/3} \left(\frac{2k-1}{2n} \pi \right)} \right)^{3/2} \\ \sigma_k &= \frac{\sigma_P}{2^{1/2} |c_k|^{2/3} \cos^{1/3} \left(\frac{2k-1}{2n} \pi \right) \left(\sum_{k=1}^{n/2} \frac{|c_k|^{2/3}}{\cos^{2/3} \left(\frac{2k-1}{2n} \pi \right)} \right)^{1/2}}. \end{aligned} \quad (\text{D.64})$$

The optimal L_0 goes as

$$L_0 \in O\left(\frac{n}{\sigma_P}\right). \quad (\text{D.65})$$

Now, the cost for estimating the interpolant with a variance of σ_P and a bias ϵ using Gaussian phase estimation is

$$C_{\text{est}} \in O\left(\frac{\log 1/\epsilon}{\sigma_P}\right) \quad (\text{D.66})$$

On the other hand, using Lemma 4, if one wishes to estimate the energy using a single ancillar qubit approach, the cost of estimating the energy is

$$C_{\text{est},1\text{-qubit}} \in O\left(\frac{\log 1/\epsilon \log \log 1/\epsilon}{\epsilon}\right), \quad (\text{D.67})$$

where ϵ is the semi-deterministic error coming from the Heisenberg-limited estimation algorithms like IQAE [43] or the one in Ref. [48], or non-Heisenberg-limited alternatives like the semi-classical QPE [49, 50], the single-qubit version of the textbook QPE [51]. \square

E Cost asymptotics for expectation values

Here we provide proofs for Lemma 18 and Theorem 19 as stated in Section 4 of the main paper.

Proof of Lemma 18. For scalar functions $f(s)$, derivatives of $\exp(f(s))$ can be expressed through the complete Bell polynomials via Faà di Bruno's formula.

$$\partial_s^n e^{f(s)} = Y_n(f'(s), f''(s), \dots, f^{(n)}(s))e^{f(s)} \quad (\text{E.1})$$

For operator exponentials such as $\exp(-i\tilde{H}_s t)$, derivatives can be expressed via repeated application of Duhamel's formula. Yet these expressions are always upper bounded by the commuting (scalar) case [52], so that

$$\left\| \partial_s^n e^{-i\tilde{H}_s t} \right\| \leq Y_n\left(t \left\| \partial_s \tilde{H}_s \right\|, t \left\| \partial_s^2 \tilde{H}_s \right\|, \dots, t \left\| \partial_s^n \tilde{H}_s \right\| \right). \quad (\text{E.2})$$

Note that the exponential disappeared in the bound since it has norm one. Applying Lemma 1 and invoking the fact that Y_n is monotonic in each argument, this is upper bounded by

$$Y_n\left(\left(2j^j (e^2 c)^{j+1}\right)_{j=1}^n\right). \quad (\text{E.3})$$

An explicit formula for this is given by

$$Y_n\left(\left(2j^j (e^2 c)^{j+1}\right)_{j=1}^n\right) = \sum_D \frac{n!}{d_1! \dots d_n!} \prod_{j=1}^n \left(\frac{2j^j (e^2 c)^{j+1}}{j!}\right)^{d_j} \quad (\text{E.4})$$

where D is a sum over all indices $(d_j)_{j=1}^n$ such that $d_j \geq 0$ and

$$\sum_{j=1}^n d_j j = n. \quad (\text{E.5})$$

Using a Stirling-type bound

$$\frac{1}{j!} \leq \left(\frac{e}{j}\right)^j \frac{1}{\sqrt{2\pi}} \quad (\text{E.6})$$

allows us to write

$$\begin{aligned} Y_n \left((2j^j (e^2 c)^{j+1})_{j=1}^n \right) &\leq \sum_D \frac{n!}{d_1! \dots d_n!} \prod_{j=1}^n \left(\sqrt{\frac{2}{\pi}} e^j (e^2 c)^{j+1} \right)^{d_j} \\ &= (e^3 c)^n \sum_D \frac{n!}{d_1! \dots d_n!} \prod_{j=1}^n \left(\sqrt{\frac{2}{\pi}} e^2 c \right)^{d_j} \\ &= (e^3 c)^n Y_n(\sqrt{2/\pi} e^2 c, \sqrt{2/\pi} e^2 c, \dots, \sqrt{2/\pi} e^2 c) \\ &= (e^3 c)^n B_n(\sqrt{2/\pi} e^2 c). \end{aligned} \quad (\text{E.7})$$

In the second line we brought out n factors of ec using the condition on the indices D , and we identified Y_n evaluated the same at every argument to be the single-variable Bell (or Touchard) polynomial B_n . We can bound the size of $B_n(\sqrt{2/\pi} e^2 c)$ [53] by

$$B_n(\sqrt{2/\pi} e^2 c) \leq \left(\frac{n}{\log(1 + \sqrt{\frac{\pi}{2}} n / (e^2 c))} \right)^n \quad (\text{E.8})$$

for all $n > 0$, with $n = 0$ defined by the limit (which is 1). With this,

$$\left\| \partial_s^n e^{-i\tilde{H}_s t} \right\| \leq \left(\frac{e^3 c n}{\log(1 + \sqrt{\frac{\pi}{2}} n / (e^2 c))} \right)^n \leq \left(\frac{e^3 c n}{2} \right)^n \left(1 + \sqrt{\frac{8}{\pi}} \frac{e^2 c}{n} \right)^n \quad (\text{E.9})$$

where we've used the bound $1/\log(1+x) \leq 1/2 + 1/x$. Again, this inequality is valid for $n = 0$ via the limit, which is always one.

With this bound on the ST formula derivatives, we now turn to bounding $\partial_s^n O_s(t)$. Applying the binomial theorem and triangle inequality to (2.13),

$$\begin{aligned} \frac{\|\partial_s^n O_s(t)\|}{\|O\|} &\leq \sum_{p=0}^n \binom{n}{p} \left\| \partial_s^p e^{it\tilde{H}_s} \right\| \left\| \partial_s^{n-p} e^{-it\tilde{H}_s} \right\| \\ &\leq \left(\frac{e^3 c}{2} \right)^n \sum_{p=0}^n \binom{n}{p} p^p (n-p)^{n-p} \left(1 + \sqrt{\frac{8}{\pi}} \frac{e^2 c}{p} \right)^p \left(1 + \sqrt{\frac{8}{\pi}} \frac{e^2 c}{n-p} \right)^{n-p}. \end{aligned} \quad (\text{E.10})$$

At this point, it will be fruitful to consider two regimes. Recall that c encodes information about the simulation time.

In the case where $c > n$, we have

$$\begin{aligned} \frac{\|\partial_s^n O_s\|}{\|O\|} &\leq \left(\frac{e^3 c}{2}\right)^n \sum_{p=0}^n \binom{n}{p} (c + \sqrt{8/\pi} e^2 c)^p (c + \sqrt{8/\pi} e^2 c)^{n-p} \\ &\leq \left(\frac{e^3 c}{2}\right)^n c^n \left(1 + \sqrt{8/\pi} e^2\right)^n \sum_{p=0}^n \binom{n}{p} \\ &= \left(c \sqrt{e^3 (1 + \sqrt{8/\pi} e^2)}\right)^{2n}. \end{aligned} \tag{E.11}$$

This implies a relative error in the polynomial fit bounded by

$$|E_{n-1}(0)| < \left(129 \frac{c^2 a}{n}\right)^n. \tag{E.12}$$

In the case where $c \leq n$, the approximation

$$\left(1 + e^2 \sqrt{\frac{8c}{\pi p}}\right)^p < e^{ce^2 \sqrt{8/\pi}} \tag{E.13}$$

holds and is not so crude. Applying this to (E.10),

$$\frac{\|\partial_s^n O_s\|}{\|O\|} \leq \left(\frac{e^3 c}{2}\right)^n n! \sum_{p=0}^n \frac{p^p (n-p)^{n-p}}{p! (n-p)!} e^{4ce^2 \sqrt{2/\pi}}. \tag{E.14}$$

Regrouping and employing a Stirling bound where appropriate,

$$\begin{aligned} \frac{\|\partial_s^n O_s\|}{\|O\|} &\leq e^{4ce^2 \sqrt{2/\pi}} \left(\frac{e^3 c}{2}\right)^n n! \left(2 \frac{e^n}{\sqrt{2\pi n}} + \frac{e^n}{2\pi} \sum_{p=1}^{n-1} \frac{1}{\sqrt{p(n-p)}}\right) \\ &\leq e^{4ce^2 \sqrt{2/\pi}} \left(\frac{e^4 c}{2}\right)^n n! \left(\frac{2}{\sqrt{2\pi n}} + \frac{\sqrt{n-1}}{2\pi}\right) \\ &\leq \frac{1}{2\pi} (\sqrt{8\pi} + \sqrt{n-1}) \left(\frac{e^4 c}{2}\right)^n n! e^{4ce^2 \sqrt{2/\pi}} \\ &\leq \sqrt{\frac{2n}{\pi}} \left(\frac{e^4 c}{2}\right)^n n! e^{4ce^2 \sqrt{2/\pi}} \end{aligned} \tag{E.15}$$

After another Stirling bound, this gives a corresponding interpolation error of

$$E_{n-1}(0) < 2\sqrt{2n} (6ca)^n e^{24c}. \tag{E.16}$$

□

Proof of Theorem 19. Let $f(s) = \langle O_s(t) \rangle / \|O\|$ be the normalized expectation value under Trotter evolution. Our interpolation algorithm produces an estimate \bar{f} of $f(0)$ which we require to be accurate within ϵ .

$$|f(0) - \bar{f}| \leq \epsilon \quad (\text{E.17})$$

There is the interpolation error from the polynomial $P_{n-1}f$ fitting f assuming perfect interpolation points $(s_i, f(s_i))$. But $f(s_i)$ can only be estimated; let's call \tilde{y}_i this estimate. The error in \tilde{y}_i in our analysis comes from the statistical error inherent in the estimation protocol as well as the error in the fractional query procedure for a $1/s$ evolution. We can independently consider the interpolation error and the data error via the triangle inequality.

$$\begin{aligned} |f(0) - \bar{f}| &\leq |f(0) - P_{n-1}f(0)| + |P_{n-1}f(0) - \tilde{P}_{n-1}f(0)| \\ &\leq \epsilon_{\text{int}} + L_n \epsilon_{\text{data}} \end{aligned} \quad (\text{E.18})$$

Here L_n is the Lebesgue constant of the interpolation, essentially a condition number, and ϵ_{data} is an upper bound on the error in the data. $\tilde{P}_{n-1}f$ is the fit to the imperfect data and $P_{n-1}f$ the fit to the perfect data $(s_i, f(s_i))$. For generic interpolation nodes, L_n can grow rapidly; however, for the set of Chebyshev nodes we obtain a near-optimal value [54].

$$L_n \leq \frac{2}{\pi} \log(n+1) + 1 \quad (\text{E.19})$$

Since we want the total error to be within a threshold ϵ , we can require

$$\epsilon_{\text{data}} \leq \frac{\epsilon}{2L_n}, \quad \epsilon_{\text{int}} \leq \frac{\epsilon}{2}. \quad (\text{E.20})$$

Given these error bounds, we can now turn to the cost of acquiring the data points. Because O/γ can be block encoded, the expectation value calculation can be encoded as an amplitude estimation problem. Specifically, a Hadamard test circuit gives the amplitude

$$\frac{1 + \langle O_{s_i}(t) \rangle / \gamma}{2}. \quad (\text{E.21})$$

If we estimate this amplitude to within accuracy $\epsilon_{\text{data}} \|O\| / 2\gamma$, we can estimate $f(s_i)$ within ϵ_{data} . Using Iterative Quantum Amplitude Estimation [43], we can obtain this estimate using a Grover iterate G constructed from two Hadamard test oracles. The number of Grover oracles N_G required is given by

$$N_G \leq \frac{200\gamma L_n}{\|O\| \epsilon_{\text{data}}} \log \left(\frac{2n}{\delta} \log_2 \left(\frac{\gamma L_n \pi}{\|O\| \epsilon_{\text{data}}} \right) \right) \quad (\text{E.22})$$

to ensure probability $1-\delta$ of all data being within ϵ_{data} of the true value. Each G requires two Hadamard tests, and each Hadamard oracle calls a (controlled) ST evolution once. The number of controlled exponentials needed for a single data point at value s_i is in

$$O\left(\frac{N_k}{|s_i|} \log 1/\epsilon_{\text{data}}\right) \quad (\text{E.23})$$

where $N_k = 2m5^{k-1}$, and where the logarithm comes from the need for fractional queries with QSVT. There is also a $O(1)$ overhead associated with the fractional queries and IQAE. Altogether, the number of exponentials for a single data point is in

$$O\left(N_G \times 2 \times \frac{N_k}{|s_i|} \log 1/\epsilon_{\text{data}}\right). \quad (\text{E.24})$$

Therefore, the total number N_{exp} of generating all $n/2$ data points (we only need half due to symmetry) is in

$$N_{\text{exp}} \in O\left(N_G N_k \sum_{i=1}^{n/2} \frac{1}{s_i} \log(1/\epsilon_{\text{data}})\right). \quad (\text{E.25})$$

Plugging in (E.22) for N_G above and summing over $1/s_i$ using Lemma 3,

$$\begin{aligned} N_{\text{exp}} &\in O\left(\frac{\gamma N_k L_n n}{\|O\| \epsilon a} (\log n) \log\left(\frac{2n}{\delta} \log_2\left(\frac{\gamma L_n \pi}{\|O\| \epsilon}\right)\right) \log(1/\epsilon_{\text{data}})\right) \\ &\subset \tilde{O}\left(\frac{n}{a\epsilon} \log 1/\delta\right) \end{aligned} \quad (\text{E.26})$$

where \tilde{O} suppresses factors logarithmic in n and ϵ . We also employed our assumption that $\gamma/\|O\| \in O(1)$. The number of nodes n and the interpolation interval $[-a, a]$ will be determined by ϵ_{int} , the interpolation error assuming perfect data. To apply our error bounds from the previous subsection, choose a to satisfy Lemma 1, i.e. $ca < \pi/20$, while also taking $1/a \in O(c)$.

Choose $n \geq \lceil c \rceil$. Then the second bound of Lemma 18 holds. From the interpolation error, we must satisfy

$$2\sqrt{2}n(6ca)^n e^{24c} < \epsilon/2 \quad (\text{E.27})$$

which in turn can be satisfied provided that

$$4\sqrt{2}n(6e^{24}ca)^n < \epsilon \quad (\text{E.28})$$

since $n \geq c$. Choose a such that $6e^{24}ca = 1/2$, which is consistent with our previous conditions on a . Then, to satisfy the error bound, n can satisfy

$$4\sqrt{2}n2^{-n} < \epsilon. \quad (\text{E.29})$$

This can be solved using the -1 branch of the LambertW function.

$$n > -\frac{\text{LambertW}_{-1}\left(-\epsilon \log 2/(4\sqrt{2})\right)}{\log 2}. \quad (\text{E.30})$$

The appropriate asymptotics is $n \in O(\log(1/\epsilon))$. By taking $n = n_*$ where

$$n_* = \max \left\{ \lceil c \rceil, \left\lceil -\frac{\text{LambertW}_{-1}\left(-\epsilon \log 2/(4\sqrt{2})\right)}{\log 2} \right\rceil \right\} \quad (\text{E.31})$$

$$\in O(\max\{c, \log(1/\epsilon)\}) \quad (\text{E.32})$$

we satisfy all required constraints and arrive at our final asymptotic scaling.

$$N_{\text{exp}} \in \tilde{O}\left(\max\{c, \log(1/\epsilon)\}c\epsilon^{-1} \log(1/\delta)\right) \quad (\text{E.33})$$

□

Remark. *In the proof above, we set $n > c$ from the beginning, in order to use the second of the two bounds from Lemma 18. Together with $1/a \in O(c)$ this condemns us to a suboptimal c^2 scaling in the large c limit. However, using the first bound instead of the second would not help us, since the nc^2/a term in that bound must be order one.*