Resource-efficient quantum matrix processing with commutator scaling

Arul Rhik Mazumder[1,2], James Watson[2], and Samson Wang[1]

[1]Caltech
[2]Carnegie Mellon University
[3]Google Quantum AI

Abstract

We develop quantum algorithms for estimating properties of general matrix functions using product formulae, with applications to phase estimation, Green's function evaluation, and sampling from time-evolved states. The resulting methods exhibit low depth, commutator scaling similar to that found for product formulae, and require only a single ancillary qubit. Our central primitive applies Richardson extrapolation to product formulae. By considering a randomized compilation scheme, we also give a protocol to statistically approximate measurement statistics of quantum states, which extends previous settings beyond observable estimation. We give refined analyses of gate complexities for k-local and power law systems; matrix ensembles with long tails; systems where commutator scaling is only understood up to a fixed order; and settings where there is a prior on input states – those with fixed Fermion number.

Contents

1 Introduction

1.1 Motivation and Background

Simulating quantum systems has long been a foundational motivation for quantum computing and remains one of the most promising paths toward achieving near-term quantum advantage. In recent years, there has been a growing focus on early fault-tolerant (EFT) quantum computing—an emerging paradigm that seeks to implement practically useful quantum algorithms on small-scale error-corrected devices, constrained by limited qubit counts and circuit depth. EFT algorithms are designed to make the most of these constraints, typically by trading off quantum resources for increased classical post-processing or runtime. Many EFT approaches build directly on time evolution as a core subroutine, making it a central building block in this regime.

A wide range of EFT algorithms leverage Hamiltonian simulation to perform tasks such as quantum phase estimation, ground state preparation, and solving linear systems. Notable examples that will relate to our work include the seminal work of Lin and Tong [LT22], who achieve Heisenberg-limited precision using just one ancilla qubit and classical post-processing (if time evolution is specified as an oracle); Wan et al. [WBC22b], who reduce circuit depth via randomized measurements; and Wang et al. [WMB24], who generalize the randomized approach to matrix function algorithms. Outside of this, the EFT paradigm has motivated a body of work on reducing quantum resources using classical overhead [Cam21; DLT22; Wan+23; Wan+25; Kis+25].

At the heart of many of these EFT algorithms lies time evolution—the simulation of unitary dynamics e^{-iHT} for a Hamiltonian H over time T. There are four main algorithmic families for simulating time evolution: product formulae (commonly referred to as Trotterization) [Llo96], linear combinations of unitaries (LCU) [CW12], quantum walks [BC12], and qubitization [LC19]. Among these, qubitization holds the best asymptotic complexities, and product formulae stand out as especially well-suited to the EFT setting. While they have worse asymptotic scaling than more advanced methods, they require no ancillary qubits or complex block-encoding (allowing them to be run on quantum devices in the present day), and their component operations decompose naturally into native gates, which simplifies compilation and reduces overhead. Importantly, they also preserve physical properties such as symmetries and locality [Tra+20], and their error bounds—while formally scaling with nested commutators—tend to perform much better empirically than worst-case predictions suggest [HHZ19].

These advantages make product formulae a compelling candidate for near-term EFT implementations. As such, understanding and improving time evolution via product formulae in the EFT regime remains a central challenge—and opportunity—for early practical quantum algorithms.

1.2 Main Contributions

We develop and refine analyses for two simple but powerful subroutines. The first enables the estimation of the following quantities:

$$\text{estimate } \mathrm{Tr}[\rho f(H)] \text{ and } \mathrm{Tr}[f(H)\rho f(H)^\dagger O] \text{ to additive precision } \varepsilon\,, \tag{1}$$

where ρ is a quantum state, O is an observable, and $f(H)$ denotes the eigenvalue transform of the function f on a Hermitian operator H. The second allows us to approximate measurement statistics of the state proportional to $f(H)\,|\psi\rangle$. Specifically, we ask

$$\text{return } \vec{v} \text{ such that } \|\vec{v} - \vec{p}\|_2 \leq \varepsilon \text{ for } p_i := |\,\langle i|\, f(H)\,|\psi\rangle\,|^2. \tag{2}$$

We view these two tasks as broadly applicable framework that accommodates a lot of quantum algorithms targeting practical applications – in both outputting scalar quantities or in sampling tasks. While some algorithmic approaches in the literature produce a quantum state as their output, most truly end-to-end use cases ultimately require a classical output [Dal+23]. To make this concrete, consider a simple example: if f is an inverse function and $\rho = |\vec{y}\rangle\langle\vec{y}|$ is a pure state that encodes a data vector $\vec{y}$, this setup corresponds to the well-known quantum linear systems problem [HHL09]. If the observable O is taken to be a local measurement, then the resulting quantity gives a marginal of the solution to the linear system. More generally, Tasks (1) and (2) can be used to extract spectral information about a target matrix H.

We give algorithms to instantiate Tasks (1) and (2) whenever given Fourier approximation to the function f, and ability to Trotterize A. The central idea is to consider an extrapolation of circuits consisting of product formulae of different step sizes, and extracting estimates of this extrapolation in a sample-efficient manner via randomization. The algorithms exhibit commutator scaling, and use a single ancillary qubit. Further, the gate complexity can be substantially improved over algorithms using product formula without extrapolation.

We remark that Task (1) is also instantiable with the same gate complexity as our approach using recent work [Cha+25], which introduces a more general primitive. Nevertheless, we still elucidate Task (1) in full using our randomized compiler, and provides a simpler streamlined proof specified for matrix function tasks. Task (2) has not traditionally been considered before in the literature aside from in [WMB24], and it opens up a new way to probe output distributions of quantum states despite being in an early fault-tolerant setting where matrix processing is not instantiated fully coherently.

The finer details of the complexity depends on the properties of the function of interest, f, and are presented in Theorems 24 and 27. Generally, both algorithms require $\widetilde{O}(1/\varepsilon^2)$ samples from quantum circuits to estimate both quantities to additive precision ε, with additional factors again depending on the function. In the following sections, we elucidate the complexity of a core subroutine and compare with prior art. We also discuss complexities of specific tasks in Section 1.5.

Our second contribution is to demonstrate refined gate complexities in a number of settings that exploit commutator scaling. These were all previously unknown for extrapolated product formulae circuits or not explicitly instantiated. We cover

- k-local systems, subsuming with it models of power-law interactions. Here, previous studies on extrapolation ([WW24a; Cha+25]) could not account for this on account the commutator factor diverging.

- Scenarios where commutator scaling for standard product formula is only understood for a fixed order. Previous studies could only exploit commutator scaling when scaling is understood at every order.

- Interpolation of our scaling with randomized formula, exploiting the ideas of [Gün+25], which can be particularly advantageous for matrix ensembles with long tails.
- Settings where a prior is known on the input state: states with well-defined Fermion number.

1.3 Warm-up: compiling a time signal

In this section we discuss a special case of Task (1): estimating the time signal $\mathrm{Tr}[\rho e^{-iHT}]$ for a given time T. This serves both as a warm-up and as essentially a core primitive through our work. Further, this gives us a chance to compare complexities with other primitive to instantiate time evolution in a transparent.

Our main algorithmic tool is to apply Richardson extrapolation to product formula approximations of time signals. The idea is simple algorithmically: simulate the system at multiple step sizes $\delta_j = s_j T$, then combine the results using a carefully chosen linear combination that cancels out low-order errors. With m such simulations, this technique can suppress errors up to order $O(\delta^{m+1})$, if step sizes lie within a radius of convergence, to be defined.

This approach is well-suited to early fault-tolerant (EFT) quantum devices. Instead of requiring deeper circuits to achieve higher accuracy, Richardson extrapolation reduces quantum depth through classical post-processing—trading shorter circuits for more samples. Furthermore, we demonstrate in this section that the sample blowup can be made very mild by use of a randomized protocol.

Let us start by introducing some notation.

Definition 1. *We consider Hermitian matrices $H \in \mathbb{C}^{2^n \times 2^n}$, expressed as a sum of Γ terms:*

$$H = \sum_{\gamma=1}^{\Gamma} H_\gamma \,,$$

where it is assumed that the unitaries $\exp(-iH_\gamma t)$ are instantiable for all t and γ in $O(1)$ gate depth. We define $\Lambda := \sum_{\gamma=1}^{\Gamma} \|H_\gamma\|$, as the sum of the strength of individual terms, where $\|\cdot\|$ denotes the operator norm.

Generically, the gate complexity of algorithms for time evolution depend on Λ. However, the complexity of p-th order product formulae can be shown to depend on a refined quantity $(\alpha_{\mathrm{comm}}^{(p+1)})^{1/p}$, which we define below.

Definition 2 (Commutator factor)**.** *Given the decomposition $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ from the previous definition, we define the order-j commutator operator*

$$\alpha_{\mathrm{comm}}^{(j)} := \sum_{\gamma_1,\ldots,\gamma_j=1}^{\Gamma} \left\| \left[H_{\gamma_1}, \left[H_{\gamma_2}, \ldots, \left[H_{\gamma_{j-1}}, H_{\gamma_j} \right] \ldots \right] \right] \right\|,$$

that is, the sum over all j-fold nested commutators formed from the Hamiltonian terms H_γ.

method	ancillary qubits	max depth per sample $\widetilde{O}(\cdot)$	sample overhead $O(\cdot)$	commutator scaling?
Qubitization [LC17]	$\lceil \log(\Gamma) + 3 \rceil$	$\Gamma\left(\Lambda T + \frac{\log(1/\varepsilon)}{\log\log(1/\varepsilon)}\right)$	$1/\varepsilon^2$	
Product formulae [Chi+21a]	1	$\Gamma(\alpha_{\text{comm}}^{(p+1)})^{1/p}\, T^{1+1/p}(1/\varepsilon)^{1/p}$	$1/\varepsilon^2$	✓
Random compiler [WBC22b]	1	$\Lambda^2 T^2$	$1/\varepsilon^2$	
Our work (i)	1	$\Gamma\,(\lambda_{\text{comm}} T)^{(1+1/p)} \log(1/\varepsilon)$	$\frac{(\log\log(1/\varepsilon))^2}{\varepsilon^2}$	✓
Our work (ii)	1	$\Gamma\left(\Lambda T + (\alpha_{\text{comm}}^{(p+1)})^{1/p} T^{1+1/p}\right) \log(1/\varepsilon)$	$\frac{(\log\log(1/\varepsilon))^2}{\varepsilon^2}$	✓
Our work (iii)	1	$\left(\Gamma_H (\lambda_{\text{comm}} T)^{1+1/p} + \Lambda_H^2 T^2\right) \log(1/\varepsilon)$	$\frac{(\log\log(1/\varepsilon))^2}{\varepsilon^2}$	✓

Table 1: **Complexity comparison for the task of approximating a time signal** $\mathrm{Tr}[\rho e^{-iHT}]$**.** In order to allow comparison with the random compiler of [WBC22b] we have assumed that the Hamiltonian is given as a linear combination of Γ Pauli matrices. The quantity $\alpha_{\text{comm}}^{(p+1)} \leq \frac{1}{2}(2\Lambda)^{p+1}$ is a sum of nested commutators, and $\lambda_{\text{comm}} \leq 4\Lambda$ is constructed from sums of nested commutators.

It is known that in many cases $(\alpha_{\text{comm}}^{(p+1)})^{1/p} << \Lambda$, which leads to significant advantage of product formula over other methods in dependence on Hamiltonian parameters.

We present the resource overhead of our algorithm framework for the compilation of a time signal in Table 1.

While qubitization achieves the best asymptotic scaling in terms of precision, we argue it is not likely the strongest candidate for early fault-tolerant (EFT) devices. This is due to the confluence of a few properties: qubitization has a larger space overhead, it requires complex multi-qubit-controlled gates, and (to our knowledge) there are not currently ways to improve its gate complexity by exploiting physical properties of the Hamiltonian such as locality. Product formulae, on the other hand, can be hardware-friendly and can take advantage of commutator scaling, often performing well in practice. However, they suffer from exponentially worse scaling with respect to the target error. The randomized algorithm proposed in [WBC22b] (built in the spirit of qDRIFT [Cam19] yet distinct and catered specifically for this task) sidesteps the exponentially worse error dependence, and is unique in having no dependence on the number of Hamiltonian terms Γ. However, this comes at the cost of forfeiting commutator scaling and incurring a quadratic scaling with evolution time. Further, we note that the approach of [WBC22b] requires a decomposition of H in the Pauli basis, which may not always be efficiently obtainable or be the most useful decomposition. In contrast, our approach works as long as H can be decomposed into efficiently simulable Hamiltonians, which is a more general data access assumption.

Our algorithm retains the advantageous properties of product formulae—such as no additional space overhead, subquadratic scaling in the time parameter T, and a commutator-sensitive error structure—while achieving additive error ε using circuits of depth that scales sub-polynomially with $1/\varepsilon$. This yields an exponential improvement in precision dependence compared to standard Trotter methods.

1.4 Commutator scaling for extrapolation

Our algorithms are found to depend on a quantity λ_{comm} which we define below.

Definition 3 (Extrapolated commutator factor)**.**

$$\lambda_{\text{comm}} := \sup_{\substack{j \in \mathbb{Z}^+ \geq \sigma_m \\ 1 \leq \ell \leq K}} \left(\sum_{\substack{j_1, \ldots, j_\ell \in \mathbb{Z}^+_{\geq p} \\ j_1 + \cdots + j_\ell = j}} \prod_{\kappa=1}^{\ell} \frac{\alpha_{\text{comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2} \right)^{\frac{1}{j+\ell}}.$$

We remark that any bound on this quantity requires understanding of $\alpha_{\text{comm}}^{(j)}$ at every order j. For instance, if $\alpha_{\text{comm}}^{(j)} = O(c^j)$ for some constant c, then $\lambda_{\text{comm}} = O((\alpha_{\text{comm}}^{(p+1)})^{(1/p)}$, which is the case of second-quantized Hamiltonians in the plane wave basis [Cha+25]. However, knowledge of $\alpha_{\text{comm}}^{(j)}$ at every order is not always available, or known bounds are too large and leads to a divergence in λ_{comm}, as is the case for k-local systems. In the rest of this section we discuss resolutions of these issues, as well as refinements to the scaling in row (i) of Table 1.

1.4.1 Partial Randomization

The most relevant early fault-tolerant alternatives to Algorithm 1 are randomized Hamiltonian simulation methods, such as those in [WBC22b]. These methods avoid explicit dependence on the number of Hamiltonian terms Γ by introducing randomness into the simulation steps, with costs that scale with the operator norm $\lambda = \Lambda$. However, when λ is dominated by a few large terms, performance can degrade.

Our approach refines this idea by combining Richardson extrapolation with partial randomization, as developed in [Gün+25]. This yields a hybrid scheme that interpolates between deterministic Trotter formulas and fully randomized strategies, enabling flexible tradeoffs between dependence on ε^{-1} and λ.

In standard Trotter methods, the cost scales linearly with Γ and favorably with ε^{-1}, but becomes impractical when $\Gamma = O(N^4)$, as in many electronic structure problems. Randomized product formulae eliminate the dependence on Γ, but their cost scales as $O(\lambda^2 t^2)$, which can be suboptimal when λ is large.

By decomposing the Hamiltonian into a few dominant terms treated deterministically and a long tail of small terms handled randomly, we balance the strengths of both approaches. Embedding this decomposition into a Richardson extrapolation framework allows systematic error suppression while keeping circuit depth low.

We now state an informal version of our main theorem, summarizing the resource complexity of estimating $\text{Tr}[Ze^{-iHT}]$ with this method:

Theorem 4 (Extrapolated partial randomization (informal version of Theorem 44))**.** *Given any decomposition $H = H_A + H + B$, the time signal $\text{Tr}[\rho e^{-iHT}]$ can be estimated to additive error ε*

and with success probability at least $(1-\delta)$ *using Algorithm 2, with resources*

$$C_{gate} = \tilde{O}\left(L_A \left(a_{\max}\, \Upsilon\, \tilde{\lambda}_{comm}\, T\right)^{1+\frac{1}{p}} + \lambda_B^2 T^2\right),$$
$$C_{sample} = \tilde{O}\left(\frac{1}{\varepsilon^2} \log\left(\frac{1}{\delta}\right)\right)$$

where L_A *is the number of deterministic (Trotterized) terms,* λ_B *is the total weight of the randomized tail terms, and* $\tilde{\lambda}_{comm} \leq \lambda_{\text{comm}}$ *depends on the norms of nested commutators.*

This result constitutes row (iii) of Table 1.

1.4.2 k-Local Systems

For k-local systems on n qubits with maximum on-site energy g, we have $(\alpha_{\text{comm}}^{(p+1)})^{1/p} = O(kgp\Lambda^{1/p})$. Our key result is that in our extrapolation algorithms (as well as prior art [WW24a; Cha+25]) we can consider $\lambda_{\text{comm}} \to \widetilde{\lambda}_{k\text{-local}}$ where for the time signal compilation

$$\widetilde{\lambda}_{k\text{-local}} = O\Big(kg(p\Lambda^{1/p} + g\log(ngT/\varepsilon))\Big).$$

and for general applications, there is an additional additive logarithmic term. Thus, the key features of commutator scaling for k-local systems are replicated, with mild logarithmic overhead. We present analysis in detail in Section 4.

1.4.3 Exploiting Fixed Fermion Number

For systems confined to an η-fermion subspace, we can improve our analysis by using the fermionic semi-norm. In this context, the operators preserve the number of fermions, meaning they map states with exactly η electrons back to states with the same number of electrons. As a result, the gate complexity depends on the fermionic semi-norms of nested commutators rather than more general norms. Although the overall bounds remain similar, they are now expressed in terms of a specialized quantity $\lambda_{\text{comm}} \to \lambda_{\text{comm}}^{(\eta)}$, which is derived from a tighter parameter $\alpha_{\text{comm}}^{(\eta)}$ that can be significantly smaller than the original α_{comm}, as elucidated in earlier studies [MCS22; SHC21; Low+23]. We present the explicit form of $\lambda_{\text{comm}}^{(\eta)}$ in Section 5.

1.5 Applications

We instantiate our algorithms for two scalar tasks: phase estimation and estimating Green functions in many-body physics. We also demonstrate an application for recovering distribution information of time-evolved states.

1.5.1 Phase Estimation

In phase estimation, the goal is to compute the ground energy E_0 of a given Hamiltonian H. We use the early-fault-tolerant approach proposed by Lin and Tong [LT22] in which we approximate the

Heaviside function using a Fourier series approximation. The key idea is the Heaviside function serves as a filter for eigenvalues, and prepares a cumulative distribution function corresponding to eigenspectrum, weighted by ρ. Thus, by probing values of the (approximate) Heaviside function, one can determine the ground state energy by finding the first jump of the cumulative distribution. It is important to note that Lin and Tong [LT22] specify their algorithm in terms of time evolution oracles, and it is necessary to instantiate it. We instantiate their oracle with our primitive, and compile the cumulative distribution function in a randomized fashion, which competes with the fully randomized algorithm of [WBC22b].

We directly apply our primitive from Algorithm 1 to implement the cumulative distribution function,

$$\widetilde{C}(x) = \mathrm{Tr}[\rho\, \widetilde{\Theta}(xI - \kappa H)]$$

where $\widetilde{\Theta}$ is a filter that approximates the Heaviside function. More details are provided in Section 7 but our main results are in the following theorem.

Theorem 5 (Phase estimation (informal version of Theorem 51))**.** *The ground state energy estimation problem can be solved to precision ε and with success probability at least $(1-\delta)$ using the above algorithm with resources*

$$C_{gate} = \tilde{O}\left(\Gamma\left(\frac{\lambda_{comm}}{\varepsilon}\right)^{(1+1/p)}\right)$$

$$C_{sample} = \tilde{O}\left(\frac{1}{\eta^2}\right)$$

where λ_{comm} is the commutator and Γ is the number of terms in the Hamiltonian decomposition used for Trotterization.

method	ancillary qubits	max depth per sample $\widetilde{O}(\cdot)$	samples $\widetilde{O}(\cdot)$	commutator scaling?
QET-U + Qubitization [DLT22; LC17]	$\lceil \log(\Gamma) \rceil + 3$	$\Gamma\Lambda/\varepsilon$	$1/\eta$	
QET-U + Trotter [DLT22]	1	$\Gamma(\alpha_{\mathrm{comm}}^{(p+1)})^{1/p}/(\eta^{1/p}\varepsilon^{1+1/p})$	$1/\eta$	✓
Random compiler [WBC22b]	1	Λ^2/ε^2	$1/\eta^2$	
Randomized extrapolation	1	$\Gamma\lambda_{\mathrm{comm}}^{1+1/p}/\varepsilon^{1+1/p}$	$1/\eta^2$	✓

Table 2: **Phase estimation complexity comparison.** For Hamiltonian $H = \sum_{\ell=1}^{\Gamma} h_\ell$, $\Lambda = \sum_{\ell=1}^{\Gamma} \|h_\ell\|$, $\eta = \mathrm{Tr}[\rho\Pi_0]$ the initial state overlap, $\lambda_{\mathrm{comm}} \leq 4\Lambda$ defined in Lemma 11. We have assumed constant order p.

We contextualize our phase estimation algorithm with other reading approaches in the literature in Table 2. We compare our algorithm with a refined early fault-tolerant approach, QET-U, where

time evolution is instantiated with qubitization [LC17] and Trotter formulae. We also compare our algorithm with the randomized compilation approach of [WBC22b]. In a similar comparison of our method with other modern approaches, we observe the following trade-offs. While QET-U + Qubitization achieves the best asymptotic circuit depth, it is not early fault-tolerant due to its logarithmic scaling in the number of ancillary qubits. QET-U + Trotter, on the other hand, retains commutator scaling and requires only a single ancilla qubit, but introduces an additional scaling overhead of $\eta^{-1/p}$. The Random Compiler method loses commutator scaling entirely and exhibits quadratic scaling with respect to ε^{-2}.

Our approach combines the most favorable aspects of these techniques: it is early fault-tolerant, preserves commutator scaling, and achieves subquadratic scaling in ε^{-1}, albeit at the cost of quadratically worse sample complexity.

1.5.2 Green's Functions

Green's function characterize how a quantum system responds to external perturbations and is essential for determining spectral properties such as excitation energies. In quantum many-body physics, the Green's function provides detailed information about the system's dynamical behavior, including how particles propagate and interact over time. It plays a central role in predicting measurable quantities like absorption spectra, density of states, and lifetimes of excited states. To obtain the Green's function, we estimate the resolvent operator, defined as

$$R(\omega + i\Gamma_{\text{broad}}, \hat{H}) = (\omega + i\Gamma_{\text{broad}} - \hat{H})^{-1},$$

where $\hat{H}$ is the system Hamiltonian and ω is the energy variable. A small positive broadening factor ($\Gamma_{\text{broad}} > 0$) is introduced to shift the poles of the resolvent into the complex plane, ensuring numerical stability and convergence. This allows for a more tractable approximation of the Green's function.

We apply Algorithm 1 to implement the Linear Combination of Unitaries $h(\omega + i\Gamma_{\text{broad}} - \hat{H})^{-1}$ approximation of the resolvent operator $R(\omega + i\Gamma_{\text{broad}}, \hat{H})$

Theorem 6 (Resolvent estimation (informal version of Theorem 55))**.** *The resolvent operator*

$$R(\omega + i\Gamma_{broad}, \hat{H}) = (\omega + i\Gamma_{broad} - \hat{H})^{-1}$$

can be estimated to additive error ε and with success probability at least $(1-\delta)$ using the randomized Richardson algorithm with resource costs

$$C_{gate} = \tilde{O}\left(\Gamma\left(\frac{\lambda_{comm}}{\Gamma_{broad}}\right)^{1+\frac{1}{p}}\right)$$

$$C_{sample} = \tilde{O}\left(\frac{1}{\Gamma_{broad}^2\,\varepsilon^2}\right)$$

where $a_{\max}$ is the maximum expansion coefficient, Υ is the truncation factor, λ_{comm} is the commutator bound, and Γ_{broad} is the broadening parameter.

1.5.3 Time-evolved states

Theorem 7 (Sampling distribution of time-evolved states (informal))**.** *We give an algorithm to approximate the measurement probability distribution* $p_i := |\langle i| e^{-iHT} |\psi\rangle|^2$ *in* ℓ_2 *norm error* ε*. The algorithm uses* $\widetilde{O}(1/\varepsilon^2)$ *samples of a circuit with one ancillary qubit and gate depth* $C_{\text{gate}} = O\left(\Gamma \left(a_{\max} \Upsilon \lambda_{\text{comm}} T\right)^{1+\frac{1}{p}} \log(1/\varepsilon)\right)$.

Acknowledgments

The authors thank the following for helpful and insightful discussions (listed in alphabetical order): Fernando Brandão, Garnet Chan, Abhinav Deshpande, Sam McArdle, Yuan Su. AM was supported by Caltech's Summer Undergraduate Research Fellowship (SURF) program. We acknowledge funding provided by the Institute for Quantum Information and Matter, an NSF Physics Frontiers Center (NSF Grant PHY-2317110).

2 Series expansion for time signal

In this section, we derive a series expansion for the time signal under the Trotter formulas. Our goal is to express the time signal as a power series in the inverse Trotter step number $1/r$. This series expansion will be the key tool used for Richardson extrapolation. We start with the variation of parameters formula:

$$e^{(A+B)t} = e^{At} + \int_0^t e^{A(t-\tau)} B e^{(A+B)\tau} d\tau, \tag{3}$$

and then use the following lemma from [WW24b, Lemma 2].

To analyze general Trotter-like methods more systematically, we consider the class of product formulae known as staged product formulae, which encompass both Lie–Trotter and higher-order Suzuki formulas.

Definition 8 (Staged Product Formula [Chi+21a])**.** *A staged product formula* $P(t)$ *is an approximation to* e^{-iHt} *of the form*

$$P(t) := \prod_{\upsilon=1}^{\Upsilon} \prod_{\gamma=1}^{\Gamma} e^{-it\, a_{(\upsilon,\gamma)}\, H_{\pi_\upsilon(\gamma)}}, \tag{4}$$

where:

- $\Upsilon \in \mathbb{Z}^+$ *is the number of* stages*,*
- $\Gamma \in \mathbb{Z}^+$ *is the number of terms per stage,*
- $a_{(\upsilon,\gamma)} \in \mathbb{R}$ *are real-valued* coefficients*,*
- $\pi_\upsilon \in S_\Gamma$ *are* permutations*,*

- *and each H_j is a term in a decomposition $H = \sum_{j=1}^{\Gamma} H_j$ of the full Hamiltonian.*

This formulation enables a structured sequence of exponentials that can be tailored through the choice of $a_{(\upsilon,\gamma)}$ and π_υ. The order of a product formula refers to how well it approximates e^{-iHt} in the small-time limit. Specifically, a formula is said to be of order $p \in \mathbb{Z}^+$ if it satisfies

$$P(t) - e^{-iHt} = O(t^{p+1}) \quad \text{as } t \to 0. \tag{5}$$

In what follows, we use this formulation to express the approximation error of such formulas as a power series in $1/r$, and thereby facilitate the use of extrapolation techniques.

2.1 Error Expansion for Time-evolved Observables

Lemma 9 (Effective Hamiltonian Error Series ([WW24b], Lemma 2))**.** *Let $\mathcal{P}$ be a staged product formula with coefficients $a_{(\upsilon,\gamma)}$ of order $p \in \mathbb{Z}_+$, and let H_{eff} be the effective Hamiltonian of $\mathcal{P}$ defined by the relation*

$$\mathcal{P}(t) = e^{-itH_{\text{eff}}(t)},$$

for $t \in \mathbb{R}$. Suppose that there exists a $J \in \mathbb{Z}_+$ and $C \in \mathbb{R}_+$ such that

$$\sup_{j \geq J} \alpha_{\text{comm}}^{(j)} (a_{\max} \Upsilon |t|)^j \leq C,$$

with $a_{\max} \coloneqq \max_{\upsilon,\gamma} |a_{(\upsilon,\gamma)}|$ and Υ as the number of stages of the staged product formula. Then the effective Hamiltonian can be written as a convergent series

$$H_{\text{eff}}(t) = H + \sum_{j=1}^{\infty} E_{j+1} t^j,$$

where

$$E_j \coloneqq \frac{(-i)^{j-1}}{j!} \sum_{\mathcal{J}} \binom{j}{j_1 \ldots j_n} \left(\prod_{i=1}^{n} a_i^{j_i} \right) \phi_j \left(H_{\gamma_1}^{\times j_1}, \ldots, X_{\gamma_n}^{\times j_n} \right),$$

and $n = \Upsilon\Gamma$ where Γ is the number of simpler Hamiltonians that H_{eff} is broken into (so $H_{\text{eff}} = \sum_{i=1}^{\Gamma} H_i$). Moreover, E_j satisfies the bound

$$\|E_j\| \leq \frac{(a_{\max} \Upsilon)^j}{j^2} \alpha_{\text{comm}}^{(j)}.$$

For convenience, we express the Magnus expansion Hamiltonian as

$$H_{\text{eff}}(t) = H + \Delta(t),$$

where $\Delta(s) = \sum_{j=1}^{\infty} E_{j+1} (sT)^j$ and H is the Hamiltonian of interest that we are trying to simulate. We apply the variation of parameters formula (Eq. (3)) on this H_{eff} expression to write:

$$e^{itH_{\text{eff}}(t)} = e^{iHt} + \int_0^t e^{i(t-\tau)H} i\Delta(s) e^{i\tau H_{\text{eff}}(t)} d\tau.$$

We use this formula to prove the following lemma.

Lemma 10. *Let $\mathcal{P}$ be a staged pth order product formula. Then*

$$\mathcal{P}(t) = e^{-itH_{\text{eff}}(t)}$$

be the approximate evolution operator for duration $T \in \mathbb{R}$ and Trotter step size $t = sT$, with $s = 0$ defined via the limit. Suppose that there exists a $J \in \mathbb{Z}_+$ and $C \in \mathbb{R}_+$ such that

$$\sup_{j \geq J} \alpha_{\text{comm}}^{(j)} \left(a_{\max} \Upsilon |sT|\right)^j \leq C,$$

with $a_{\max} \coloneqq \max_{\upsilon,\gamma} |a_{(\upsilon,\gamma)}|$. Let $\sigma = 2$ if $\mathcal{P}$ is symmetric, 1 otherwise. Then for any $K \in \mathbb{Z}_+$, the approximation error in $\mathcal{P}^{1/s}(sT)$ compared with the exact evolution may be expressed as

$$\text{Tr}[\rho(\mathcal{P}^{1/s}(sT) - e^{-iHT})] = \sum_{j \in \sigma\mathbb{Z}_+ \geq p} s^j \text{Tr}[\rho \tilde{E}_{j+1,K}(T)] + \text{Tr}[\rho \tilde{F}_K(T,s)].$$

Here, $\tilde{E}_{j+1,K}(T)$ and $\tilde{F}_K(T,s)$ are operators whose spectral norm $\|\cdot\|$ is bounded as

$$\begin{aligned}\|\tilde{E}_{j+1,K}(T)\| &\leq (a_{\max}\Upsilon T)^j \sum_{l=1}^{\min\{K-1,\lfloor j/p \rfloor\}} \frac{(a_{\max}\Upsilon T)^l}{l!} \sum_{\substack{j_1 \ldots j_l \in \sigma\mathbb{Z}_+ \geq p \\ j_1+\cdots+j_l=j}} \left(\prod_{\kappa=1}^{l} \frac{\alpha_{\text{comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2}\right) \\ &= \sum_{l=1}^{\min\{K-1,\lfloor j/p \rfloor\}} \frac{(a_{\max}\Upsilon T \lambda_{j,l})^{j+l}}{l!}.\end{aligned}$$

Note that

$$\alpha_{comm}^{(j_\kappa+1)} = \sum_{\gamma_1,\gamma_2 \ldots \gamma_{j_k}=1}^{\Gamma} \left[H_{\gamma_1}, \left[H_{\gamma_2}, \ldots, \left[H_{\gamma_{j_{k-1}}}, H_{\gamma_{j_k}}\right] \ldots\right]\right]$$

Where we have denoted

$$\lambda_{j,l} \coloneqq \left(\sum_{\substack{j_1 \ldots j_1 \in \sigma\mathbb{Z}_+ \geq p \\ j_1+\cdots+j_l=j}} \prod_{\kappa=1}^{l} \frac{\alpha_{\text{comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2} \right)^{1/(j+l)}.$$

Proof. Recall from [WW24b, Lemma 3] that a pth order staged product formula has error operator $\Delta(s) = \sum_{j \in \sigma\mathbb{Z}_+ \geq p} E_{j+1} \cdot (sT)^j$ wher $\sigma = 2$ if $\mathcal{P}$ is symmetric, 1 otherwise. Using the variation-of-parameters formula, we can iteratively substitute the same formula into itself multiple times. Doing so $K \in \mathbb{Z}^+$ times, we can write the following:

$$e^{iTH_{\text{eff}}(s)} = e^{iHT} + \tag{6}$$

$$+ \sum_{l=1}^{K-1} \int_0^T d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{l-1}} d\tau_l e^{i(T-\tau_1)H} i\Delta(s) e^{i(\tau_1-\tau_2)H} i\Delta(s) \ldots e^{i(\tau_{l-1}-\tau_l)H} i\Delta(s) e^{iH\tau_l} \tag{7}$$

$$+ \int_0^T d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{p-1}} d\tau_K e^{i(T-\tau_1)H} i\Delta(s) e^{i(\tau_1-\tau_2)H} i\Delta(s) \ldots e^{i(\tau_{K-1}-\tau_K)H} i\Delta(s) e^{i\tau_K H_{\text{eff}}(s)}. \tag{8}$$

Taking line (7) and expanding the definition of $\Delta(s)$,

$$\sum_{l=1}^{K-1}\int_0^T d\tau_1\int_0^{\tau_1}d\tau_2\cdots\int_0^{\tau_{l-1}}d\tau_l e^{i(T-\tau_1)H}i\Delta(s)e^{i(\tau_1-\tau_2)H}i\Delta(s)\dots e^{i\tau_{l-1}-\tau_l)H}i\Delta(s)e^{iH\tau_l}$$

$$=\sum_{l=1}^{K-1}\int_0^T d\tau_1\int_0^{\tau_1}d\tau_2\cdots\int_0^{\tau_{l-1}}d\tau_l\left(\prod_{\kappa=l}^{1}\left(\sum_{j_\kappa\in\sigma\mathbb{Z}_+\geq p}e^{i(\tau_{\kappa-1}-\tau_\kappa)H}iE_{j_\kappa+1}t^{j_\kappa}\right)\right)e^{iH\tau_l}$$

$$=\sum_{l=1}^{K-1}\int_0^T d\tau_1\int_0^{\tau_1}d\tau_2\cdots\int_0^{\tau_{l-1}}d\tau_l\left(\sum_{j\in\sigma\mathbb{Z}_+\geq pl}t^j\sum_{\substack{j_1\dots j_l\in\sigma\mathbb{Z}_+\geq p\\ j_1+\cdots+j_l=j}}\left(\prod_{\kappa=l}^{1}e^{i(\tau_{\kappa-1}-\tau_\kappa)H}iE_{j_\kappa+1}\right)\right)e^{iH\tau_l},$$

where we have denoted $\tau_0=T$. We now reinsert $t=sT$, and make a change of variables $s_i=\tau_i/T$. This gives

$$\sum_{l=1}^{K-1}T^l\int_0^1 ds_1\int_0^{s_1}ds_2\cdots\int_0^{s_{l-1}}ds_l\left(\sum_{j\in\sigma\mathbb{Z}_+\geq pl}(sT)^j\sum_{\substack{j_1\dots j_l\in\sigma\mathbb{Z}_+\geq p\\ j_1+\cdots+j_l=j}}\left(\prod_{\kappa=l}^{1}e^{i(s_{\kappa-1}-s_\kappa)TH}iE_{j_\kappa+1}\right)\right)e^{iHs_lT}.$$

Next, we regroup the sum according to the degree of s, which yields

$$\sum_{j\in\sigma\mathbb{Z}_+\geq p}(sT)^j\sum_{l=1}^{\min\{K-1,\lfloor j/p\rfloor\}}T^l\int_0^1 ds_1\int_0^{s_1}ds_2\cdots\int_0^{s_{l-1}}ds_l\sum_{\substack{j_1\dots j_l\in\sigma\mathbb{Z}_+\geq p\\ j_1+\cdots+j_l=j}}\left(\prod_{\kappa=l}^{1}e^{iT(s_{\kappa-1}-s_\kappa)H}iE_{j_\kappa+1}\right)e^{iHs_lT}$$

$$=\sum_{j\in\sigma\mathbb{Z}_+\geq p}s^j\tilde{E}_{j+1,K}(T).$$

Here, we have defined

$$\tilde{E}_{j+1,K}(T)\coloneqq$$
$$\sum_{l=1}^{\min\{K-1,\lfloor j/p\rfloor\}}T^{j+l}\int_0^1 ds_1\int_0^{s_1}ds_2\cdots\int_0^{s_{l-1}}ds_l\sum_{\substack{j_1\dots j_l\in\sigma\mathbb{Z}_+\geq p\\ j_1+\cdots+j_l=j}}\left(\prod_{\kappa=l}^{1}e^{i(s_{\kappa-1}-s_\kappa)TH}iE_{j_\kappa+1}\right)e^{is_lTH}. \quad (9)$$

We now want to put bounds on the norm of $\tilde{E}_{j+1,K}$. Using the triangle inequality, unitarity of $e^{i\tau H}$, and evaluating the remaining integral,

$$\|\tilde{E}_{j+1,K}(T)\|\leq\sum_{l=1}^{\min\{K-1,\lfloor j/p\rfloor\}}T^{j+l}\int_0^1\int_0^{s_1}\cdots\int_0^{s_{l-1}}ds_l\sum_{\substack{j_1\dots j_l\in\sigma\mathbb{Z}_+\geq p\\ j_1+\cdots+j_l=j}}\left(\prod_{\kappa=1}^{l}\|E_{j_\kappa+1}\|\right)$$

$$\leq T^j\sum_{l=1}^{\min\{K-1,\lfloor j/p\rfloor\}}\frac{T^l}{l!}\sum_{\substack{j_1\dots j_l\in\sigma\mathbb{Z}_+\geq p\\ j_1+\cdots+j_l=j}}\left(\prod_{\kappa=1}^{l}\|E_{j_\kappa+1}\|\right).$$

Applying Lemma 9,

$$\|\tilde{E}_{j+1,K}(T)\| \leq T^j \sum_{l=1}^{\min\{K-1,\lfloor j/p\rfloor\}} \frac{T^l}{l!} \sum_{\substack{j_1\ldots j_l\in\sigma\mathbb{Z}_+\geq p\\ j_1+\cdots+j_l=j}} \left(\prod_{\kappa=1}^{l} \alpha_{\text{comm}}^{(j_\kappa+1)} \frac{(a_{\max}\Upsilon)^{j_\kappa+1}}{(j_\kappa+1)^2}\right)$$
$$= (a_{\max}\Upsilon T)^j \sum_{l=1}^{\min\{K-1,\lfloor j/p\rfloor\}} \frac{(a_{\max}\Upsilon T)^l}{l!} \sum_{\substack{j_1\ldots j_l\in\sigma\mathbb{Z}_+\geq p\\ j_1+\cdots+j_l=j}} \left(\prod_{\kappa=1}^{l} \frac{\alpha_{\text{comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2}\right).$$

So far we have considered the terms in line (7). We now consider line (8), which we will denote as $\tilde{F}_K(T,s)$. Applying the triangle inequality and utilising unitarity in a similar manner as above, we can check that the operator norm of $\tilde{F}_K(T,s)$ is bounded by

$$\begin{aligned}\|\tilde{F}_K(T,s)\| &\leq \frac{T^K}{K!}\|\Delta(s)\|^K \\ &\leq \frac{T^K}{K!}\left(\sum_{j\in\sigma\mathbb{Z}_+\geq p} \|E_{j+1}\|(sT)^j\right)^K \end{aligned} \tag{10}$$

which vanishes for $K\to\infty$, so we take this limit. □

2.2 Extrapolation Error Bound

In this subsection we characterize the extrapolation error, assuming exact computation of Trotterized time series. We first state a lemma which bounds the extrapolation error for any Richardson extrapolation scheme for any Hamiltonian assuming the conditions of Lemma 10.

Lemma 11 (Generic Richardson extrapolation error for time signal)**.** *Let $\mathcal{P}$ be a staged pth order product formula of symmetry class σ as Definition 8. For a target evolution time T let*

$$\mathcal{P}_{p,m}^{(R)}(T) := \sum_{k=1}^{m} b_k \mathcal{P}^{1/s_k}(s_k T)\,,$$

denote an m-term Richardson extrapolation, with ascending sequence of Trotter steps $r_k = 1/s_k \in \mathbb{Z}_+$, which cancel the powers $s^\sigma, s^{\sigma 2}, \ldots, s^{\sigma(m-1)}$. Then, the error in the extrapolation of a Trotterized time signal, as compared to an exact time signal, satisfies

$$\left|\text{Tr}[\rho(\mathcal{P}_{p,m}^{(R)}(T) - e^{-iHT})]\right| \leq \|\vec{b}\|_1 \sum_{\substack{j\in\sigma\mathbb{Z}_+\\ j\geq\sigma m}} s_1^j \sum_{l=1}^{K} \frac{(a_{\max}\Upsilon\lambda_{\text{comm}}T_{j,l})^{j+l}}{l!}\,.$$

for any quantum state ρ, where $\|\vec{b}\|_1 = \sum_k |b_k|$, $K = \left\lceil \frac{\sigma m}{p} \right\rceil$, and we denote

$$\lambda_{\text{comm}} := \sup_{\substack{j\in\sigma\mathbb{Z}_+\geq\sigma m\\ 1\leq l\leq K}} \lambda_{j,l}\,, \qquad \lambda_{j,l} := \left(\sum_{\substack{j_1\ldots j_1\in\sigma\mathbb{Z}_+\geq p\\ j_1+\cdots+j_l=j}} \prod_{\kappa=1}^{l} \frac{\alpha_{\text{comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2}\right)^{1/(j+l)}.$$

Proof. From Lemma 10, recall we have the error series in s as

$$\mathrm{Tr}[\rho(\mathcal{P}^{1/s}(sT) - e^{-iHT})] = \sum_{j\in\sigma\mathbb{Z}_+\geq p} s^j \mathrm{Tr}[\rho\tilde{E}_{j+1,K}(T)] + \mathrm{Tr}[\rho\tilde{F}_K(T,s)]\,,$$

for any $K \in \mathbb{Z}_+$, with $\|\tilde{F}_K(T,s)\| = O(s^{Kp})$. Richardson extrapolation with m terms removes all terms of size up to and including $O(s^{\sigma(m-1)})$ in the series. In particular, if we choose K such that $Kp > \sigma(m-1)$ then the m-term Richardson series only cancels terms in $\tilde{E}$. Concretely, choose $K = \lceil \sigma m/p \rceil$. We denote the the Richardson extrapolation of $\mathcal{P}^{1/s}(sT)$ as $\mathcal{P}^{(R)}_{p,m}(T) := \sum_{k=1}^m b_k \mathcal{P}^{1/s_k}(s_k T)$, where b_k are the extrapolation coefficients, and write the extrapolation error as

$$\mathrm{Tr}[\rho(\mathcal{P}^{(R)}_{p,m}(T) - e^{-iHT})] = \sum_{k=1}^m b_k \|R_{\sigma(m-1)}(T,s_k)\|.$$

where $R_{\sigma(m-1)}(T,s_k)$ denotes the Taylor remainder of degree $\sigma(m-1)$ satisfying

$$R_{\sigma(m-1)}(T,s_k) := \sum_{\substack{j\in\sigma\mathbb{Z}_+ \\ j\geq\sigma m}} s_k^j \tilde{E}_{j+1,K}(T) + \tilde{F}_K(T,s_k)\,,$$

for each inverse Trotter step s_k. The Taylor remainder for inverse Trotter step s_k has size

$$\begin{aligned}
\left\|R_{\sigma(m-1)}(T,s_k)\right\| &\leq \sum_{\substack{j\in\sigma\mathbb{Z}_+ \\ j\geq\sigma m}} s_k^j \left\|\tilde{E}_{j+1,K}(T)\right\| + \left\|\tilde{F}_K(T,s_k)\right\| \\
&\leq \sum_{\substack{j\in\sigma\mathbb{Z}_+ \\ j\geq\sigma m}} s_k^j \sum_{l=1}^{K-1} \frac{\left(a_{\max}\Upsilon\lambda_{\mathrm{comm}}T_{j,l}\right)^{j+l}}{l!} + \sum_{\substack{j\in\sigma Z_+ \\ j\geq Kp}} s_k^j \frac{\left(a_{\max}\Upsilon\lambda_{\mathrm{comm}}T_{j,l}\right)^{j+K}}{K!} \\
&\leq \sum_{\substack{j\in\sigma\mathbb{Z}_+ \\ j\geq\sigma m}} s_k^j \sum_{l=1}^{K} \frac{\left(a_{\max}\Upsilon\lambda_{\mathrm{comm}}T_{j,l}\right)^{j+l}}{l!}\,,
\end{aligned}$$

where we have used Lemma 10, denoted $\lambda_{j,l}$ as in the statement of this Lemma, and in the final inequality have extended the sum in the second term to start at the smaller value $j \geq \sigma m$ and merged the sum over l. We remark that in the final bound the second term is wholly independent of s_k, and thus the largest bound of remainder terms comes at inverse step number s_1. Using this fact and Hölder's inequality, the size of the error can be bounded as

$$\left|\mathrm{Tr}[\rho(\mathcal{P}^{(R)}_{p,m}(T) - e^{-iHT})]\right| \leq \|\vec{b}\|_1 \cdot \max_k \left\|R_{\sigma(m-1)}(T,s_k)\right\| \leq \|\vec{b}\|_1 \sum_{\substack{j\in\sigma\mathbb{Z}_+ \\ j\geq\sigma m}} s_1^j \sum_{l=1}^{K} \frac{\left(a_{\max}\Upsilon\lambda_{\mathrm{comm}}T_{j,l}\right)^{j+l}}{l!}\,.$$

where we have denoted $b := (b_1, .., b_m)$ as the vector of extrapolation coefficients. □

We then simplify the Richardson extrapolation error bound by adding a lower bound on the number of Trotter steps required.

Lemma 12. *Adopt the setting of Lemma 11, and choose the largest inverse Trotter step number s_1 such that $s_1 a_{\max} \Upsilon \lambda_{\mathrm{comm}} T < 1/2$. Define $\eta := \max\{1, a_{\max} \Upsilon T \lambda_{\mathrm{comm}}\}$. The m-term Richardson extrapolation error for a staged pth-order product formula of symmetry class σ is bounded as*

$$\left|\mathrm{Tr}[\rho(\mathcal{P}_{p,m}^{(R)}(T) - e^{-iHT})]\right| \leq 4\|\vec{b}\|_1 \, \eta^{\lceil \sigma m/p \rceil} \, (s_1 a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{\sigma m} \ .$$

Proof. We continue from Eq. (2.2) of Lemma 11. Note that

$$\begin{aligned} \sum_{l=1}^{K} \frac{(a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{j+l}}{l!} &\leq (a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{j} \sum_{l=1}^{K} \frac{(a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{l}}{l!} \\ &\leq (a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{j} \, \eta^K (e-1) \, . \end{aligned}$$

Thus, the sum over j in Eq. (2.2) can be bounded

$$\begin{aligned} \sum_{\substack{j \in \sigma\mathbb{Z}_+ \\ j \geq \sigma m}} s_1^j \sum_{l=1}^{K} \frac{(a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{j+l}}{l!} &\leq \eta^K (e-1) \sum_{\substack{j \in \sigma\mathbb{Z}_+ \\ j \geq \sigma m}} s_1^j \, (a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{j} \\ &\leq \eta^K (e-1) \, (s_1 a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{\sigma m} \sum_{\substack{j \in \sigma\mathbb{Z}_+ \\ j \geq 0}} s_1^j \, (a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{j} \\ &\leq \eta^K (e-1) \, (s_1 a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{\sigma m} \sum_{\substack{j \in \sigma\mathbb{Z}_+ \\ j \geq 0}} \left(\frac{1}{2}\right)^{j} \\ &\leq 4\eta^K \, (s_1 a_{\max} \Upsilon \lambda_{\mathrm{comm}} T)^{\sigma m} \ . \end{aligned}$$

□

2.3 Well-conditioned Extrapolation Strategy

We now adopt the extrapolation strategy from [LKW19] to obtain an approximation with error $O(\log(m) s^{2m}) = O(s^m)$, using coefficients whose condition number grows only as $O(\log m)$.

Lemma 13. *(Well-conditioned Richardson extrapolation [LKW19] (Lemma 5 in [WW24b])) Let $f \in C^{2m+2}([-1,1])$ be an even, real-valued function of s, and let P_j and R_j be the degree j Taylor polynomial and Taylor remainder, respectively, such that $f(s) = P_j(s) + R_j(s)$. Let*

$$F^{(m)}(s) = \sum_{k=1}^{m} b_k f(s_k)$$

be the unique Richardson extrapolation of f at points $s_1, s_2, \ldots s_m$ given by

$$s_k = \frac{s}{q_k}\,; \qquad q_k = q_{\text{scale}} \left\lceil \frac{\sqrt{8}m/\pi}{\sin(\pi(2k-1)/8m)} \right\rceil, \quad k \in \{1, \ldots, m\}\,,$$

where $q_{\text{scale}} \in \mathbb{Z}_+$ satisfies $m \leq q_k/q_{\text{scale}} \leq 3m^2$ [LKW19], and b_k given by

$$b_k = \prod_{i \neq k} \frac{x_i^2}{x_i^2 - x_k^2}\,.$$

Then, the extrapolated function satisfies

$$F^{(m)}(s) = f(0) + \sum_{k=1}^{m} b_k R_{2m}(s_k)\,,$$

and $\|\vec{b}\|_1 = O(\log m)$.

With this better defined extrapolation strategy, we can further refined our error bound as provided in the following lemma.

Lemma 14 (Richardson extrapolation error with well-conditioned extrapolation)**.** *Denote the extrapolation error using Richardson extrapolation as ε_R as*

$$\varepsilon_R := \left| \text{Tr}[\rho e^{-iHT}] - \text{Tr}[\rho\, \mathcal{P}_{p,m}^{(R)}(T)] \right|\,.$$

We suppose that $\mathcal{P}_{p,m}^{(R)}(T)$ is constructed using the well-conditioned extrapolation strategy of Lemma 13. With this, to obtain extrapolation error $\varepsilon_R \leq \varepsilon$ it is sufficient to run $m = O(\log(1/\varepsilon))$ different extrapolation circuits, each with a maximum of $O\Big((a_{\max}\Upsilon\lambda_{\text{comm}}T)^{(1+1/p)} \log(1/\varepsilon)\Big)$ Trotter steps.

Proof. For the extrapolation error, recall Lemma 12 gives

$$\varepsilon_R \leq 4\|\vec{b}\|_1\, \eta^{\lceil \sigma m/p \rceil}\, (s_1 a_{\max}\Upsilon\lambda_{\text{comm}}T)^{\sigma m}\,,$$

where $\eta := \max\{1, a_{\max}\Upsilon T\lambda_{\text{comm}}\}$. Using the extrapolation strategy of Lemma 13 with a large enough r_{scale} we can ensure that $a_{\max}\Upsilon\lambda_{\text{comm}}T > 1$. Similarly we can use this to set ε such that we have the bound

$$\varepsilon_R \leq 4\|\vec{b}\|_1\, s_1^{\sigma m}\, (a_{\max}\Upsilon\lambda_{\text{comm}}T)^{\sigma m + \lceil \sigma m/p \rceil} \leq \varepsilon,$$

which implies that a sufficient base Trotter step number is anything satisfying

$$r_1 = 1/s_1 \geq \left\lceil (a_{\max}\Upsilon\lambda_{\text{comm}}T)^{1+\frac{1}{\sigma m}\left\lceil \frac{\sigma m}{p} \right\rceil} \left(\frac{4\|\vec{b}\|_1}{\varepsilon} \right)^{\frac{1}{\sigma m}} \right\rceil\,.$$

As $(1/\varepsilon)^{1/\log(1/\varepsilon)} = O(1)$ we can take $m = O(\log(1/\varepsilon))$

$$r_1 = O\left((a_{\max}\Upsilon\lambda_{\text{comm}}T)^{(1+1/p)}\right)$$

and thus the largest number of Trotter steps satisfies

$$r_k \leq 3mr_1 = O\left((a_{\max}\Upsilon\lambda_{\text{comm}}T)^{(1+1/p)}\log(1/\varepsilon)\right).$$

□

3 Randomized algorithm framework for matrix functions

We now present and analyze a randomized algorithm to estimate $\text{Tr}[Zf(A)]$, where Z is an operator with bounded Schatten 1-norm $\|Z\|_1$. In many cases, it is natural to take $Z = \rho$, a quantum state. This algorithm forms a central building block for estimating more complex quantities such as $\text{Tr}[f(A)\rho f(A)^\dagger O]$, which arise in applications our applications of Phase Estimation and Green's function estimation.

3.1 Generalized Error Series and Error for Bounded Observables

Before introducing the main algorithm, we generalize the results from Lemmas 10 and 14 to a more general operator Z instead of ρ. Therefore we get the following lemmas:

Lemma 15 (Observable error expansion for staged product formulae)**.** *Let $\mathcal{P}$ be a staged pth-order product formula of symmetry class $\sigma \in \{1, 2\}$, and suppose the effective Hamiltonian expansion assumptions of Lemma 10 hold. Let $A = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be the Hamiltonian of interest, and let Z be an operator. Then for any $s \in \mathbb{R}$ and any integer $K \geq 1$, the observable expectation value satisfies the expansion*

$$\text{Tr}\left[Z\,\mathcal{P}^{1/s}(sT)\right] = \text{Tr}\left[Z\,e^{-iAT}\right] + \sum_{j\in\sigma\mathbb{Z}_+\geq p} s^j\,\text{Tr}\left[Z\,\tilde{E}_{j+1,K}(T)\right] + \text{Tr}\left[Z\,\tilde{F}_K(T,s)\right],$$

where the operators $\tilde{E}_{j+1,K}(T)$ and $\tilde{F}_K(T,s)$ as defined in Lemma 15.

Note that Lemma 15 follow very directly from the proof of Lemma 10, since this original proof is independent of the properties of the operator. Similarly, we also derive the following corresponding lemma to bound the number of Trotter steps.

Lemma 16 (Richardson extrapolation error for general observables)**.** *Let Z be an operator of an $O(1)$ Schatten 1-norm $\|Z\|_1$ and suppose that $\mathcal{P}^{(R)}_{p,m}(T)$ is an m-term Richardson extrapolation constructed from a staged pth-order product formula $\mathcal{P}$ of symmetry class σ, using the well-conditioned extrapolation strategy of Lemma 13. Denote the extrapolation error as*

$$\varepsilon_R := \left|\text{Tr}\left[Z\,e^{-iHT}\right] - \text{Tr}\left[Z\,\mathcal{P}^{(R)}_{p,m}(T)\right]\right|.$$

Then, to obtain error $\varepsilon_R \leq \varepsilon$, it is sufficient to run $m = O(\log(1/\varepsilon))$ extrapolation circuits, each using at most

$$O\Big((a_{\max}\,\Upsilon\,\lambda_{\text{comm}}\,T)^{1+\frac{1}{p}}\log\left(1/\varepsilon\right)\Big)$$

Trotter steps.

Proof. The proof parallels that of Lemma 14, replacing the quantum state ρ with a bounded operator Z in the trace expression. The key observation is that all bounds in Lemmas 10 and 11 depend on the operator norm of the effective Hamiltonian series and not on properties specific to ρ. Since Z is bounded and appears linearly in the trace, from Lemma 15 we get that:

$$\text{Tr}\left[Z\,\mathcal{P}^{1/s}(sT)\right] = \text{Tr}\left[Z\,e^{-iHT}\right] + \sum_{j\in\sigma\mathbb{Z}_+\geq p} s^j\,\text{Tr}\left[Z\,\tilde{E}_{j+1,K}(T)\right] + \text{Tr}\left[Z\,\tilde{F}_K(T,s)\right].$$

Applying the same analysis as in Lemma 12

$$\begin{aligned}
\varepsilon_R &= \left|\text{Tr}\left[Z\,\mathcal{P}^{1/s}(sT)\right] - \text{Tr}\left[Ze^{-iHT}\right]\right| \leq \varepsilon \\
&\Rightarrow \left|\text{Tr}\left[Z\left(\mathcal{P}^{1/s}(sT) - e^{-iHT}\right)\right]\right| \leq \varepsilon \\
&\Rightarrow \|Z\|_1 \cdot \left\|\mathcal{P}^{1/s}(sT) - e^{-iHT}\right\|_\infty \leq \varepsilon \qquad \text{(Hölder's inequality)} \\
&\Rightarrow 4\|Z\|_1\|\vec{b}\|_1\,\eta^{\lceil\sigma m/p\rceil}\,(s_1 a_{\max}\Upsilon\lambda_{\text{comm}}T)^{\sigma m} \leq \varepsilon
\end{aligned}$$

With $\eta := \max\{1, a_{\max}\Upsilon\lambda_{\text{comm}}T\}$. Using the extrapolation strategy of Lemma 13 with large enough r_{scale} we can ensure we can fix a long-time regime such that $a_{\max}\Upsilon\lambda_{\text{comm}}T > 1$ thus $\eta = a_{\max}\Upsilon\lambda_{\text{comm}}T$.

Now given this expression and considering that $\|\vec{b}\|_1 = O(\log m)$ and $m = O(\log(1/\varepsilon))$, we find to bound $\varepsilon_R \leq \varepsilon$:

$$\begin{aligned}
4\|Z\|_1\|\vec{b}\|s_1^{\sigma m}(a_{\max}\Upsilon\lambda_{\text{comm}}T)^{\lceil\sigma m/p\rceil+\sigma m} &\leq \varepsilon \\
\Rightarrow \frac{1}{s_1^{\sigma m}} &\geq \frac{4\|Z\|_1\|\vec{b}\|(a_{\max}\Upsilon\lambda_{\text{comm}}T)^{\lceil\sigma m/p\rceil+\sigma m}}{\varepsilon} \\
\Rightarrow r_1 = \frac{1}{s_1} &\geq \left(\frac{4\|Z\|_1\|\vec{b}\|}{\varepsilon}\right)^{\frac{1}{\sigma m}}(a_{\max}\Upsilon\lambda_{\text{comm}}T)^{1+\frac{1}{\sigma m}\lceil\frac{\sigma m}{p}\rceil}
\end{aligned}$$

Since we consider $\|Z\|_1 = O(1)$, $m = O(\log(1/\varepsilon))$, and $\|\vec{b}\| = O(\log m) = O(\log\log(1/\varepsilon))$, then $\frac{1}{\varepsilon}^{1/(\log\varepsilon)} = O(1)$, we get that as the base Trotter step count:

$$r_1 = O\Big((a_{\max}\Upsilon\lambda_{\text{comm}}T)^{1+1/p}\Big)$$

And thus the largest number of Trotter steps satisfies

$$r_k \leq 3mr_1 = O\Big((a_{\max}\Upsilon\lambda_{\text{comm}}T)^{1+1/p}\log(1/\varepsilon)\Big)$$

□

Remark 17. *Note that in general* $x^{\frac{1}{k}} = \exp\left(\frac{\ln x}{k}\right)$*, so we can write*

$$\begin{aligned}
\left(\frac{4\|Z\|_1\|\vec{b}\|}{\varepsilon_R}\right)^{\frac{1}{\sigma m}} &= \exp\left(\frac{\ln(4\|Z\|_1\|\vec{b}\|) + \ln(1/\varepsilon_R)}{\sigma m}\right) \\
&= \exp\left(\frac{\ln(4\|Z\|_1\|\vec{b}\|)}{\sigma m} + \frac{1}{c\sigma}\right) \qquad \textit{since } m = c\log(1/\varepsilon)) \\
&= \exp\left(\frac{\ln(4\|Z\|_1) + \ln\|\vec{b}\|}{\sigma m} + \frac{1}{c\sigma}\right) \\
&= \exp\left(\frac{\ln(4\|Z\|_1)}{\sigma m} + \frac{O(\log\log m)}{\sigma m} + \frac{1}{c\sigma}\right)
\end{aligned}$$

Thus the step complexity remains the same as long as $\|Z\|_1 = O(\log(1/\varepsilon))$

Remark 17 will be useful in Lemma 26.

3.2 Richardson Approximation of properties of $f(A)$

To motivate our algorithm, we first describe how to approximate properties of $f(A)$ using Richardson extrapolation combined with a truncated Fourier series.

Let $A = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be a Hermitian operator (e.g., a Hamiltonian), and consider a function $f : \mathbb{R} \to \mathbb{C}$. We approximate $f(A)$ by truncating its Fourier series expansion to K terms:

$$f(A) \approx f_K(A) := \sum_{k=1}^{K} c_k e^{iAt_k}, \tag{11}$$

where the coefficients $c_k \in \mathbb{C}$ and time parameters $t_k \in \mathbb{R}$ depend on the desired approximation accuracy, specifically on the Fourier truncation error

$$\varepsilon_{\mathrm{F}} := \left|\mathrm{Tr}[Zf(A)] - \mathrm{Tr}[Zf_K(A)]\right|.$$

We denote the maximal frequency by

$$T := \max_{1\le k\le K} t_k,$$

and the vector of Fourier coefficients as $\vec{c} = (c_1, \ldots, c_K)$ with ℓ_1-norm

$$c := \|\vec{c}\|_1 = \sum_{k=1}^{K} |c_k|.$$

Each exponential e^{iAt_k} is then approximated using an m-term Richardson extrapolation based on a p-th order product formula $\mathcal{P}$:

$$e^{iAt_k} \approx \sum_{j=1}^{m} b_j\, \mathcal{P}^{1/s_j}(s_j t_k),$$

where the coefficients b_j come from a well-conditioned Richardson extrapolation scheme, and the s_j are scaling parameters (e.g., step sizes). We define the fully discretized Fourier–Richardson approximation as

$$f_{K,m}(A) := \sum_{k=1}^{K} \sum_{j=1}^{m} c_k b_j \mathcal{P}^{1/s_j}(s_j t_k). \tag{12}$$

For convenience, denote

$$\mathcal{Z} := \sum_{k=1}^{K} \sum_{j=1}^{m} |c_k b_j|.$$

Since the coefficients factorize as

$$\mathcal{Z} = \left(\sum_{k=1}^{K} |c_k| \right) \left(\sum_{j=1}^{m} |b_j| \right) = O(c \log m) \tag{13}$$

where the $O(\log m)$ factor arises from the growth of the Richardson coefficients $\{b_j\}$ (see Lemma 13), we obtain a clean bound on the normalization constant $\mathcal{Z}$. This normalization plays a crucial role in constructing probability distributions for randomized sampling in our main algorithm.

3.3 Main Algorithmic Primitive

We now present our main randomized algorithm for estimating quantities of the form $\mathrm{Tr}[Zf(A)]$. Rather than implementing a linear combination of quantum circuits directly, we estimate the target quantity by collecting samples from a family of circuits, chosen according to a carefully designed classical probability distribution. Each circuit corresponds to a single term in the expansion of $f(A)$, and we use Hadamard test circuits to collect real and imaginary components of individual trace contributions. The final estimate is obtained by averaging over many independent samples. The exact analysis of our algorithm is provided in Theorem 23

Algorithm 1 Randomized Algorithm for Estimating $\mathrm{Tr}[Zf(A)]$

1: **Classical Preprocessing:** Compute the probability distribution:

$$\left\{\frac{|c_k b_j|}{\mathcal{Z}}\right\} \quad \text{for } k = 1, \ldots, K;\ j = 1, \ldots, M$$

2: **for** $i = 1$ to C_{sample} **do**
3: Sample an index pair (k', j') from the classical distribution
4: Prepare the Hadamard test circuit corresponding to:
- $X_{\text{Re}}^{(k')} = \text{Re}\left[\text{Tr}\left(Z\mathcal{P}^{1/s_{j'}}(s_{j'}t_{k'})\right)\right]$
- $X_{\text{Im}}^{(k')} = \text{Im}\left[\text{Tr}\left(Z\mathcal{P}^{1/s_{j'}}(s_{j'}t_{k'})\right)\right]$

5: Collect one measurement outcome (a single-shot sample) from each circuit
6: Let $X^{(k')} = X_{\text{Re}}^{(k')} + iX_{\text{Im}}^{(k')}$ be the resulting complex-valued sample
7: Multiply $X^{(k')}$ by $\|c_{k'}b_{j'}\|_1 \cdot \text{sign}(c_{k'}b_{j'})$ and store the result
8: **end for**
9: **Return:** $\widehat{\mu}_M$, the sample average over all C_{sample} values

Note that step 4 of Algorithm 1 may not always be physically possible depending on the form of operator Z. In this work, we specialize to two choices for the operator Z in the above algorithm. The first case is when Z corresponds directly to a quantum state, which allows straightforward preparation of the necessary circuits via Hadamard tests. In this case the analysis follows quite cleanly as now only $\|Z\|_1 = 1$.

The second case arises when Z can be decomposed as $Z = OU\rho$, where U is a unitary operator, O is an observable, and ρ is a quantum state (in this case our Trotter circuit $\mathcal{P}$ corresponds to U_2). We describe the circuit preparation for Step 4 of Algorithm 1 in the following subsection.

3.3.1 Explicit Circuit Preparation

In this subsection we describe how to prepare Step 4 of Algorithm 1, when $Z = OU\rho$, where U is an implementable unitary operator, O is an observable, and ρ is a quantum state. This decomposition enables the construction of the Hadamard test circuits by leveraging known methods for preparing and measuring these components. We show and prove the constructions of the circuits in the following lemmas.

Lemma 18 (Estimating $\text{Re}\left(\text{Tr}[OU_1\rho U_2^\dagger]\right)$)**.** *Let ρ be a quantum state, and let U_1, U_2, O be unitary or bounded operators with $O = O^\dagger$. Then, the expectation value*

$$\text{Re}\left(\text{Tr}[OU_1\rho U_2^\dagger]\right)$$

can be estimated using the quantum circuit that prepares the real part of the trace by:

- *Initializing the ancilla in* $|+\rangle$ *with a Hadamard gate,*
- *Applying a controlled-*U_1 *(when control is* $|1\rangle$*),*
- *Applying an anti-controlled-*U_2 *(when control is* $|0\rangle$*),*
- *Applying the observable* $X \otimes O$ *and measuring its expectation.*

Proof. We can trace the circuit after each gate to verify its outcome:

$$
\begin{aligned}
|+\rangle\langle+| \otimes \rho &= \frac{1}{2}\left(|0\rangle\langle 0| \otimes \rho + |1\rangle\langle 1| \otimes \rho + |0\rangle\langle 1| \otimes \rho + |1\rangle\langle 0| \otimes \rho\right) && \text{(initial state)} \\
&\mapsto \frac{1}{2}\left(|0\rangle\langle 0| \otimes \rho + |1\rangle\langle 1| \otimes U_1\rho U_1^\dagger + |0\rangle\langle 1| \otimes \rho U_1^\dagger + |1\rangle\langle 0| \otimes U_1\rho\right) \\
& && \text{(apply controlled-}U_1\text{)} \\
&\mapsto \frac{1}{2}\left(|0\rangle\langle 0| \otimes U_2\rho U_2^\dagger + |1\rangle\langle 1| \otimes U_1\rho U_1^\dagger + |0\rangle\langle 1| \otimes U_2\rho U_1^\dagger + |1\rangle\langle 0| \otimes U_1\rho U_2^\dagger\right) \\
& && \text{(apply anti-controlled-}U_2\text{)} \\
&\mapsto \mathrm{Tr}\left[(X \otimes O) \cdot \frac{1}{2}\left(|0\rangle\langle 1| \otimes U_2\rho U_1^\dagger + |1\rangle\langle 0| \otimes U_1\rho U_2^\dagger\right)\right] \\
& && \text{(apply } X \otimes O \text{ and take trace)} \\
&= \frac{1}{2}\left(\mathrm{Tr}[OU_2\rho U_1^\dagger] + \mathrm{Tr}[OU_1\rho U_2^\dagger]\right) && (O \text{ Hermitian}) \\
&= \mathrm{Re}\left(\mathrm{Tr}[OU_2\rho U_1^\dagger]\right) \\
&= \mathrm{Re}\left(\mathrm{Tr}[OU_1\rho U_2^\dagger]\right) && \text{(By cyclicity of trace)}
\end{aligned}
$$

□

Lemma 19 (Estimating $\mathrm{Im}\left(\mathrm{Tr}[OU_1\rho U_2^\dagger]\right)$)**.** *Let* ρ *be a quantum state, and let* U_1, U_2, O *be unitary or bounded operators with* $O = O^\dagger$*. Then, the expectation value*

$$\mathrm{Im}\left(\mathrm{Tr}[OU_1\rho U_2^\dagger]\right)$$

can be estimated using the following quantum circuit procedure:

- *Prepare an ancilla qubit in the* $|+i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i\,|1\rangle)$ *state (e.g., by applying S then H to* $|0\rangle$*),*
- *Apply a controlled-*U_1 *with the ancilla as control,*
- *Apply an anti-controlled-*U_2*,*
- *Measure the observable* $Y \otimes O$*, where* Y *acts on the ancilla.*

Proof. We can trace the circuit after each gate to verify its outcome:

$$
\begin{aligned}
|+i\rangle\langle+i| \otimes \rho &= \frac{1}{2}\left(|0\rangle\langle 0| \otimes \rho + |1\rangle\langle 1| \otimes \rho + i\,|0\rangle\langle 1| \otimes \rho - i\,|1\rangle\langle 0| \otimes \rho\right) && \text{(initial state)} \\
&\mapsto \frac{1}{2}\left(|0\rangle\langle 0| \otimes \rho + |1\rangle\langle 1| \otimes U_1\rho U_1^\dagger + i\,|0\rangle\langle 1| \otimes \rho U_1^\dagger - i\,|1\rangle\langle 0| \otimes U_1\rho\right) \\
& && \text{(apply controlled-}U_1\text{)} \\
&\mapsto \frac{1}{2}\left(|0\rangle\langle 0| \otimes U_2\rho U_2^\dagger + |1\rangle\langle 1| \otimes U_1\rho U_1^\dagger + i\,|0\rangle\langle 1| \otimes U_2\rho U_1^\dagger - i\,|1\rangle\langle 0| \otimes U_1\rho U_2^\dagger\right) \\
& && \text{(apply anti-controlled-}U_2\text{)} \\
&\mapsto \mathrm{Tr}\left[(X \otimes O) \cdot \frac{1}{2}\left(i\,|0\rangle\langle 1| \otimes U_2\rho U_1^\dagger - i\,|1\rangle\langle 0| \otimes U_1\rho U_2^\dagger\right)\right] \\
& && \text{(apply } X \otimes O \text{ and take trace)} \\
&= \frac{i}{2}\left(\mathrm{Tr}[OU_2\rho U_1^\dagger] - \mathrm{Tr}[OU_1\rho U_2^\dagger]\right) && (O \text{ Hermitian}) \\
&= \mathrm{Im}\left(\mathrm{Tr}[OU_2\rho U_1^\dagger]\right) \\
&= \mathrm{Im}\left(\mathrm{Tr}[OU_1\rho U_2^\dagger]\right) && \text{(By cyclicity of trace)}
\end{aligned}
$$

□

These circuit constructions are necessary and parallel our main results.

3.3.2 Correctness

We now show that the algorithm in Algorithm 1 is unbiased. That is, the expected value of the estimator over the sampling distribution equals the target quantity $\mathrm{Tr}[Z f_{K,m}(A)]$.

Lemma 20. *Let* (k, j) *be sampled with probability* $\Pr(k, j) = \frac{|c_k b_j|}{\mathcal{Z}}$*, and define*

$$\widehat{\mu}_M := \mathcal{Z} \cdot \mathrm{sign}(c_k b_j) \cdot \mathrm{Tr}\left[Z\mathcal{P}^{1/s_j}(s_j t_k)\right].$$

Then $\widehat{\mu}_M$ *is an unbiased estimator of* $\mathrm{Tr}[Z f_{K,m}(A)]$*, i.e.,*

$$\mathbb{E}[\widehat{\mu}_M] = \mathrm{Tr}\left[Z f_{K,m}(A)\right].$$

Proof. Note that as discussing in Equation 12

$$\mathrm{Tr}[Z f_{K,m}(A)] = \sum_{k=1}^{K}\sum_{j=1}^{m} c_k b_j \,\mathrm{Tr}[Z\mathcal{P}^{1/s_j}(s_j t_k)]$$

We verify that $\widehat{\mu}_M$ is unbiased:

$$\begin{aligned}\mathbb{E}[\widehat{\mu}_M] &= \sum_{k,j} \Pr(k,j) \cdot \mathcal{Z} \cdot \operatorname{sign}(c_k b_j) \cdot \operatorname{Tr}\left[Z\mathcal{P}^{1/s_j}(s_j t_k)\right] \\ &= \sum_{k,j} \frac{|c_k b_j|}{\mathcal{Z}} \cdot \mathcal{Z} \cdot \operatorname{sign}(c_k b_j) \cdot \operatorname{Tr}\left[Z\mathcal{P}^{1/s_j}(s_j t_k)\right] \\ &= \sum_{k,j} c_k b_j \operatorname{Tr}\left[Z\mathcal{P}^{1/s_j}(s_j t_k)\right].\end{aligned}$$

□

3.3.3 Sample Complexity

We now provide a rigorous bound on the sample complexity C_{sample} required to estimate $\operatorname{Tr}[Z f_{K,m}(A)]$ to within additive error ε with high probability.

Lemma 21. *Let $\widehat{\mu}_M$ be the unbiased estimator defined in Lemma 20, and let $\widehat{\mu}_M := \frac{1}{M}\sum_{\ell=1}^{M} X^{(\ell)}$ denote the empirical mean of M independent samples. Then, for any confidence level $\delta > 0$, with probability at least $1-\delta$, the sampling error $\varepsilon_{S[M]}$ satisfies*

$$|\widehat{\mu}_M - \mathbb{E}[\widehat{\mu}_M]| := \varepsilon_{S[M]} \le \varepsilon \quad \textit{provided} \quad M \ge \frac{(2\|Z\|_1\mathcal{Z})^2}{\varepsilon^2} \cdot \log\left(\frac{2}{\delta}\right).$$

Then to estimate to sampling error $\varepsilon_{S[M]} \le \varepsilon$, the sample complexity scales as

$$C_{\text{sample}} = O\left(\frac{\|Z\|_1^2 c^2 (\log\log(1/\varepsilon))^2}{\varepsilon^2} \cdot \log\left(\frac{1}{\delta}\right)\right).$$

Proof. Each sample $X^{(\ell)}$ lies in the interval $[-\|Z\|_1\mathcal{Z}, \|Z\|_1\mathcal{Z}]$. By Hoeffding's inequality, for any $\varepsilon > 0$,

$$\mathbb{P}\left[|\widehat{\mu}_M - \mathbb{E}[\widehat{\mu}_M]| \ge \varepsilon\right] \le 2\exp\left(-\frac{2M\varepsilon^2}{(2\|Z\|_1\mathcal{Z})^2}\right).$$

To ensure that this probability is at most δ, it suffices to choose

$$M \ge \frac{(2\|Z\|_1\mathcal{Z})^2}{\varepsilon^2} \cdot \log\left(\frac{2}{\delta}\right).$$

From the normalization factor bound (see equation (13)),

$$\mathcal{Z} = \sum_{k,j} |c_k b_j| = O(c \log m),$$

and since $m = O(\log(1/\varepsilon))$, we obtain

$$\mathcal{Z} = O(c \log\log(1/\varepsilon)).$$

Substituting into the bound for M, the sample complexity becomes

$$C_{\text{sample}} = O\left(\frac{\|Z\|_1^2 c^2 (\log\log(1/\varepsilon))^2}{\varepsilon^2} \cdot \log\left(\frac{1}{\delta}\right)\right),$$

which establishes the claim. □

3.3.4 Gate Complexity

We now analyze the gate complexity of the Algorithm 1, focusing on bounding the number of Trotter steps for each instance of $\mathcal{P}^{1/s_{j'}}(s_{j'}t_{k'})$. The following result follows from Lemma 16.

Lemma 22. *Under the assumptions of Lemma 16, to ensure the total Richardson error satisfies $\varepsilon_{\text{R}} \leq \varepsilon$, it suffices to use:*

- $m = O(\log(1/\varepsilon))$ *Richardson extrapolation points, and*
- *a maximum Trotter step count of*

$$r_{\max} = O\left(\log(c/\varepsilon) \cdot (a_{\max}\Upsilon\lambda_{\text{comm}}T)^{1+\frac{1}{p}}\right).$$

This yields a total gate complexity of

$$C_{\text{gate}} = O\left(\Gamma \cdot \log(c/\varepsilon) \cdot (a_{\max}\Upsilon\lambda_{\text{comm}}T)^{1+\frac{1}{p}}\right).$$

Proof. We consider the error due to approximating $f_K(A)$ by its Richardson-extrapolated product formula:

$$f_K(A) = \sum_{k=1}^{K} c_k e^{iAt_k} \quad \text{(from equation 11)},$$

$$f_{K,m}(A) = \sum_{k=1}^{K}\sum_{j=1}^{m} c_k b_j \mathcal{P}^{1/s_j}(s_j t_k) \quad \text{(from equation 12)}.$$

Then the total Richardson error is

$$\varepsilon_{\text{R}} := \left|\text{Tr}[Zf_K(A)] - \text{Tr}[Zf_{K,m}(A)]\right| = \left|\sum_{k=1}^{K} c_k \left(\text{Tr}[Ze^{-iAt_k}] - \sum_{j=1}^{m} b_j \,\text{Tr}\left[Z\mathcal{P}^{1/s_j}(s_j t_k)\right]\right)\right|.$$

Using the triangle inequality:

$$\varepsilon_{\text{R}} \leq \sum_{k=1}^{K} |c_k| \cdot \varepsilon_{\text{R}'}(t_k) \leq c \cdot \varepsilon_{\text{R}'}(T),$$

where $\varepsilon_{\mathrm{R}'}(t_k)$ denotes the Richardson error in approximating $\mathrm{Tr}[Ze^{-iAt_k}]$ by $\sum_{j=1}^{m} b_j \mathrm{Tr}[Z\mathcal{P}^{1/s_j}(s_j t_k)]$, and where $T := \max_k |t_k|$. To ensure the total error satisfies $\varepsilon_{\mathrm{R}} \leq \varepsilon$, it suffices to set

$$\varepsilon_{\mathrm{R}'}(T) \leq \varepsilon/c.$$

By Lemma 16, this can be achieved using

$$r_{\max} = O\left(\log(c/\varepsilon) \cdot (a_{\max}\Upsilon\lambda_{\mathrm{comm}}T)^{1+\frac{1}{p}}\right)$$

Trotter steps. Since each Trotter step applies to all Γ terms in the decomposition of A, the gate complexity is

$$C_{\mathrm{gate}} = O\left(\Gamma \cdot \log(c/\varepsilon) \cdot (a_{\max}\Upsilon\lambda_{\mathrm{comm}}T)^{1+\frac{1}{p}}\right).$$

□

3.3.5 Final Complexities

We now summarize the overall resource requirements—gate, sample, and classical preprocessing complexities—for our randomized algorithm to estimate $\mathrm{Tr}[Zf(A)]$.

Theorem 23 (Sample and Gate Complexity of Randomized Trace Estimator)**.** *To estimate* $\mathrm{Tr}[Zf(A)]$ *with error at most* ε *and success probability at least* $1-\delta$*, the randomized Richardson-extrapolated Algorithm 1 requires the following resources:*

- ***Gate complexity (per sample):***

$$C_{\mathrm{gate}} = O\left(\Gamma \cdot \log\left(\frac{c(\varepsilon/3)}{\varepsilon}\right) \cdot (a_{\max}\Upsilon\lambda_{\mathrm{comm}}\, T(\varepsilon/3))^{1+\frac{1}{p}}\right),$$

  where $T := \max_k |t_k|$*, and* $c = \|\vec{c}\|_1$*. Both* c *and* T *depend on the Fourier approximation error* $\varepsilon_{\mathrm{F}} := \varepsilon/3$*.*

- ***Sample complexity:***

$$C_{\mathrm{sample}} = O\left(\frac{\|Z\|_1^2 \cdot c(\varepsilon/3)^2 \cdot (\log\log(1/\varepsilon))^2}{\varepsilon^2} \cdot \log\left(\frac{1}{\delta}\right)\right),$$

  where $\|Z\|_1$ *is the Schatten 1-norm (i.e., trace norm) of* Z*, and again* $c = \|\vec{c}\|_1$ *is determined by the Fourier truncation at error* $\varepsilon/3$*.*

- ***Classical preprocessing time:***

$$C_{\mathrm{pre}} = O(K + \log(1/\varepsilon)),$$

  which accounts for computing the Fourier coefficients $\{c_k\}$*, the Richardson coefficients* $\{b_j\}$*, and related approximation parameters.*

Proof. Let $\widehat{\mu}_M = \frac{1}{M}\sum_{\ell=1}^{M} X^{(\ell)}$ be the empirical mean of the unbiased estimator defined in Lemma 20. We can bound the total error as

$$|\mathrm{Tr}[Zf(A)] - \widehat{\mu}_M| \leq \varepsilon_{\mathrm{F}} + \varepsilon_{\mathrm{R}} + \varepsilon_{S[M]} \leq \varepsilon$$

where:

- $\varepsilon_{\mathrm{F}} := |\mathrm{Tr}[Zf(A)] - \mathrm{Tr}[Zf_K(A)]|$,
- $\varepsilon_{\mathrm{R}} := \left|\mathrm{Tr}[Zf_K(A)] - \mathrm{Tr}[Zf_{K,m}(A)]\right|$,
- $\varepsilon_{S[M]} := \left|\mathrm{Tr}[Zf_{K,m}(A)] - \widehat{\mu}_M\right|$

We set each $\varepsilon_{\text{Fourier}} = \varepsilon_{\text{Trotter+Richardson}} = \varepsilon_{\text{Sampling},M} \leq \varepsilon/3$ and therefore substituting $\varepsilon := \varepsilon/3$ into the derived gate and sample complexities from Lemma 22 and Lemma 21, we get

- **Gate complexity (per sample):**

$$C_{\text{gate}} = O\left(\Gamma \cdot \log\left(\frac{c(\varepsilon/3)}{\varepsilon}\right) \cdot \left(a_{\max}\Upsilon\lambda_{\text{comm}}\, T(\varepsilon/3)\right)^{1+\frac{1}{p}}\right)$$

- **Sample complexity:**

$$C_{\text{sample}} = O\left(\frac{\|Z\|_1^2 \cdot c(\varepsilon/3)^2 \cdot (\log\log(1/\varepsilon))^2}{\varepsilon^2} \cdot \log\left(\frac{1}{\delta}\right)\right).$$

□

3.4 Core algorithms

Let us now examine our two primary objectives: estimating the quantities $\mathrm{Tr}[\rho f(A)]$ and $\mathrm{Tr}[f(A)\rho f(A)^\dagger O]$. We begin with the first case, $\mathrm{Tr}[\rho f(A)]$.

3.4.1 Estimating $\mathrm{Tr}[\rho f(A)]$

In this setting, the gate complexity per sample remains unchanged from the general case. However, the sample complexity simplifies due to the fact that $\|Z\|_1 = \|\rho\|_1 = 1$. As a result, we obtain:

$$O\left(\frac{c(\varepsilon/3)^2\,(\log\log(1/\varepsilon))^2}{\varepsilon^2}\right),$$

where $c = \|\vec{c}\|_1$ arises from the Fourier approximation of f to accuracy $\varepsilon/3$. This leads directly to the following result.

Theorem 24 (Sample and Gate Complexity of Algorithm 1 for Estimating $\mathrm{Tr}[\rho f(A)]$)**.** *To estimate* $\mathrm{Tr}[\rho f(A)]$ *to within error* ε*, with success probability at least* $1-\delta$*, the randomized Richardson-extrapolated Algorithm 1 requires the following resources:*

- ***Gate complexity per sample:***

$$C_{\text{gate}} = O\left(\Gamma \cdot \log\left(\frac{c(\varepsilon/3)}{\varepsilon}\right) \cdot \left(a_{\max}\Upsilon\lambda_{\text{comm}}\, T(\varepsilon/3)\right)^{1+\frac{1}{p}}\right),$$

where $T := \max_k |t_k|$ denotes the maximum absolute value of the Fourier time parameters, and $c = \|\vec{c}\|_1$ depends on the truncation error $\varepsilon_{\text{F}} := \varepsilon/3$.

- ***Sample complexity:***

$$C_{\text{sample}} = O\left(\frac{c(\varepsilon/3)^2 \cdot (\log\log(1/\varepsilon))^2}{\varepsilon^2} \cdot \log\left(\frac{1}{\delta}\right)\right),$$

exploiting the fact that $\|Z\|_1 = 1$, since $Z = \rho$ is a quantum state.

- ***Classical preprocessing time:***

$$C_{\text{pre}} = O(K + \log(1/\varepsilon)),$$

which accounts for computing the Fourier coefficients $\{c_k\}$, Richardson extrapolation coefficients $\{b_j\}$, and other relevant approximation parameters.

3.4.2 Estimating $\mathbf{Tr}[f(A)\rho f(A)^\dagger O]$

Our second estimation task—computing $\text{Tr}[f(A)\rho f(A)^\dagger O]$—is more involved than the first. We begin by recalling the Fourier expansions from equation (11):

$$f(A) \approx f_K(A) = \sum_{k=1}^{K} c_k e^{iAt_k}, \qquad f(A)^\dagger \approx f_K(A)^\dagger = \sum_{l=1}^{K} c_l^* e^{-iAt_l}.$$

Using these approximations, the target quantity becomes

$$\text{Tr}[f(A)\rho f(A)^\dagger O] \approx \sum_{k,l=1}^{K} c_k c_l^* \,\text{Tr}\left[e^{iAt_k}\rho e^{-iAt_l} O\right].$$

As in equation (12), each exponential of the form e^{iAt} is approximated using an m-term Richardson extrapolation of a p-th order product formula $\mathcal{P}$. This yields the following approximation:

$$\text{Tr}[f(A)\rho f(A)^\dagger O] \approx \sum_{k,l=1}^{K}\sum_{j,r=1}^{m} c_k c_l^* b_j b_r X_{k,l,j,r},$$

where

$$X_{k,l,j,r} := \text{Tr}\left[\mathcal{P}^{1/s_j}(s_j t_k)\rho\left(\mathcal{P}^{1/s_r}(s_r t_l)\right)^\dagger O\right].$$

This decomposition allows us to analyze the sample complexity, as stated below.

Lemma 25 (Statistical Error for Estimating $\mathrm{Tr}[f(A)\rho f(A)^\dagger O]$ using Algorithm 1)**.** *Suppose* $\|O\| = 1$. *Then, to estimate*

$$\mathrm{Tr}[f(A)\rho f(A)^\dagger O]$$

to additive error ε *with failure probability at most* δ*, the number of samples required is bounded by*

$$C_{\text{sample}} = O\left(\frac{c^4 \cdot (\log\log(1/\varepsilon))^4}{\varepsilon^2} \log\left(\frac{1}{\delta}\right)\right),$$

where $c = \|\vec{c}\|_1 = \sum_k |c_k|$*, and we assume* $\|\vec{b}\|_1 \le \log m = O(\log\log(1/\varepsilon))$.

Proof. Since $\|O\| = 1$, each term $X_{k,l,j,r}$ is a bounded observable with outcomes in $[-1, 1]$. We perform importance sampling over the index tuple (k, l, j, r), with sampling probability proportional to

$$p_{k,l,j,r} \propto |c_k c_l^* b_j b_r|.$$

The variance of the corresponding unbiased estimator is then bounded using the Cauchy–Schwarz inequality:

$$\left(\sum_{k,l,j,r} |c_k c_l^* b_j b_r|\right)^2 \le \left(\sum_k |c_k|\right)^2 \left(\sum_l |c_l|\right)^2 \left(\sum_j |b_j|\right)^2 \left(\sum_r |b_r|\right)^2.$$

Letting $c = \|\vec{c}\|_1$ and using the assumption $\|\vec{b}\|_1 \le \log m = O(\log\log(1/\varepsilon))$, we obtain a variance upper bound of

$$O\left(c^4 (\log\log(1/\varepsilon))^4\right).$$

Applying Hoeffding's inequality then yields the stated sample complexity:

$$C_{\text{sample}} = O\left(\frac{c^4 (\log\log(1/\varepsilon))^4}{\varepsilon^2} \log\left(\frac{1}{\delta}\right)\right),$$

completing the proof. □

As shown in equation (3.4.2), estimating $\mathrm{Tr}[f(A)\rho f(A)^\dagger O]$ reduces to estimating terms of the form $\mathrm{Tr}[e^{iAt_k}\rho e^{-iAt_l} O]$. To evaluate such terms efficiently, we apply Richardson extrapolation as described in equation (3.2), yielding the approximation:

$$\mathrm{Tr}[e^{iAt_k}\rho e^{-iAt_l} O] \approx \sum_{j=1}^m \sum_{r=1}^m b_j b_r \,\mathrm{Tr}\left[\mathcal{P}^{1/s_j}(s_j t_k)\rho \left(\mathcal{P}^{1/s_r}(s_r t_l)\right)^\dagger O\right],$$

where $\mathcal{P}$ is a p-th order product formula and $\{b_j\}$ are the Richardson coefficients. This expansion allows us to bound the gate complexity required to estimate $\mathrm{Tr}[f(A)\rho f(A)^\dagger O]$ to within a target error.

Lemma 26 (Richardson Extrapolation Error for $\mathrm{Tr}[f(A)\rho f(A)^\dagger O]$)**.** *Let ε_R denote the total Richardson extrapolation error:*

$$\varepsilon_R := \left| \mathrm{Tr}[e^{iAt_k}\rho e^{-iAt_l}O] - \sum_{j=1}^{m}\sum_{r=1}^{m} b_j b_r \,\mathrm{Tr}\left[\mathcal{P}^{1/s_j}(s_j t_k)\rho\left(\mathcal{P}^{1/s_r}(s_r t_l)\right)^\dagger O\right]\right|.$$

Then $\mathrm{Tr}[e^{iAt_k}\rho e^{-iAt_l}O]$ can be estimated with error $\varepsilon_R \le \varepsilon$ using gate complexity

$$C_{\mathrm{gate}} = O\left(\Gamma \cdot \log(1/\varepsilon) \cdot (a_{\max}\Upsilon\lambda_{\mathrm{comm}}\, t_{\max})^{1+\frac{1}{p}}\right),$$

where $t_{\max} := \max\{|t_k|, |t_l|\}$, and the constants are inherited from the product formula and Hamiltonian simulation setting.

Proof. We start from the definition of ε_R:

$$\begin{aligned}
\varepsilon_R &= \left| \mathrm{Tr}\left[e^{iAt_k}\rho e^{-iAt_l}O\right] - \sum_{j=1}^{m}\sum_{r=1}^{m} b_j b_r \,\mathrm{Tr}\left[\mathcal{P}^{1/s_j}(s_j t_k)\rho\left(\mathcal{P}^{1/s_r}(s_r t_l)\right)^\dagger O\right]\right| \\
&= \left| \mathrm{Tr}\left[O e^{iAt_k}\rho\left(e^{-iAt_l} - \sum_{r=1}^{m} b_r \mathcal{P}^{1/s_r}(s_r t_l)\right)\right]\right. \\
&\qquad \left. + \mathrm{Tr}\left[\left(e^{iAt_k} - \sum_{j=1}^{m} b_j \mathcal{P}^{1/s_j}(s_j t_k)\right)\rho\left(\sum_{r=1}^{m} b_r \left(\mathcal{P}^{1/s_r}(s_r t_l)\right)^\dagger O\right)\right]\right|.
\end{aligned}$$

Substituting appropriately with Z_1 and Z_2 so we can call Algorithm 1 and applying triangle inequality:

$$\varepsilon_R \le \left|\mathrm{Tr}\left[Z_1 \cdot \left(e^{-iAt_l} - \sum_{r=1}^{m} b_r \mathcal{P}^{1/s_r}(s_r t_l)\right)\right]\right| + \left|\mathrm{Tr}\left[\left(\sum_{j=1}^{m} b_j \mathcal{P}^{1/s_j}(s_j t_k) - e^{iAt_k}\right) Z_2\right]\right|,$$

where

$$Z_1 := e^{iAt_k}\rho O, \qquad Z_2 := \rho \cdot \sum_{r=1}^{m} b_r \left(\mathcal{P}^{1/s_r}(s_r t_l)\right)^\dagger O.$$

We now bound the trace norms:

- $\|Z_1\|_1 \le \|e^{iAt_k}\| \cdot \|\rho\|_1 \cdot \|O\| \le 1$, assuming ρ is a quantum state and $\|O\| \le 1$.
- For Z_2, we apply Hölder's inequality:

$$\begin{aligned}
\|Z_2\|_1 &= \left\|\rho \cdot \sum_{r=1}^{m} b_r \left(\mathcal{P}^{1/s_r}(s_r t_l)\right)^\dagger O\right\|_1 \\
&\le \|\rho\|_1 \cdot \left\|\sum_{r=1}^{m} b_r \left(\mathcal{P}^{1/s_r}(s_r t_l)\right)^\dagger O\right\| \\
&\le \sum_{r=1}^{m} |b_r| \cdot \|O\| \le \|\vec{b}\|_1 \cdot \|O\| = O(\log m).
\end{aligned}$$

Finally, note that while the norm $\|Z_2\|_1$ introduces a logarithmic dependence on m, this dependence is mild. Specifically, from Remark 17, we have $\log m = \log\log(1/\varepsilon) \leq \log(1/\varepsilon)$. Therefore, we may still invoke Lemma 16 to bound the gate depth associated with simulating each product formula segment. As a result, the total gate complexity required to estimate $\mathrm{Tr}[e^{iAt_k}\rho e^{-iAt_l}O]$ to error at most ε is

$$C_{\mathrm{gate}} = O\left(\Gamma \cdot \log(1/\varepsilon) \cdot (a_{\max}\Upsilon\lambda_{\mathrm{comm}}\, t_{\max})^{1+\frac{1}{p}}\right),$$

as claimed. □

We now consolidate the previous results into a final theorem that characterizes the sample and gate complexities for estimating $\mathrm{Tr}[f(A)\rho f(A)^\dagger O]$ using Algorithm 1. The structure of the proof follows directly from the same approach used in Theorem 23.

Theorem 27 (Sample and Gate Complexity of Algorithm 1 for Estimating $\mathrm{Tr}[f(A)\rho f(A)^\dagger O]$)**.** *To estimate* $\mathrm{Tr}[f(A)\rho f(A)^\dagger O]$ *to additive error at most* ε *and success probability at least* $1-\delta$*, the randomized Richardson-extrapolated algorithm requires the following resources:*

- ***Gate complexity (per sample):***

$$C_{\mathrm{gate}} = O\left(\Gamma \cdot \log\left(\frac{c(\varepsilon/3)}{\varepsilon}\right) \cdot (a_{\max}\Upsilon\lambda_{\mathrm{comm}}\, T(\varepsilon/3))^{1+\frac{1}{p}}\right),$$

  where $T(\varepsilon/3) := \max_k |t_k|$ *is the largest time parameter in the Fourier approximation with error* $\varepsilon/3$*, and* $c(\varepsilon/3) := \|\vec{c}\|_1$ *is the associated* ℓ_1*-norm of the Fourier coefficients.*

- ***Sample complexity:***

$$C_{\mathrm{sample}} = O\left(\frac{c(\varepsilon/3)^4 \cdot (\log\log(1/\varepsilon))^4}{\varepsilon^2} \cdot \log\left(\frac{1}{\delta}\right)\right).$$

- ***Classical preprocessing time:***

$$C_{\mathrm{pre}} = O(K + \log(1/\varepsilon)),$$

  accounting for computing the Richardson coefficients $\{b_j\}$*, Fourier coefficients* $\{c_k\}$*.*

Proof. Let $\widehat{\mu}_M := \frac{1}{M}\sum_{\ell=1}^{M}\widehat{X}^{(\ell)}$ denote the empirical mean of the unbiased estimator from Algorithm 1. We denote the total error in the estimate is $\left|\mathrm{Tr}[f(A)\rho f(A)^\dagger O] - \widehat{\mu}_M\right|$.

As before, we can decompose this into three contributions (Fourier, Richardson, and sampling):

$$\begin{aligned}\left|\mathrm{Tr}[f(A)\rho f(A)^\dagger O] - \widehat{\mu}_M\right| \leq\ & \underbrace{\left|\mathrm{Tr}[f(A)\rho f(A)^\dagger O] - \mathrm{Tr}[f_K(A)\rho f_K(A)^\dagger O]\right|}_{\varepsilon_{\mathrm{F}}} \\ & + \underbrace{\left|\mathrm{Tr}[f_K(A)\rho f_K(A)^\dagger O] - \mathrm{Tr}[f_{K,m}(A)\rho f_{K,m}(A)^\dagger O]\right|}_{\varepsilon_{\mathrm{R}}} \\ & + \underbrace{\left|\mathrm{Tr}[f_{K,m}(A)\rho f_{K,m}(A)^\dagger O] - \widehat{\mu}_M\right|}_{\varepsilon_{\mathrm{S}}}.\end{aligned}$$

To ensure the total error is bounded by ε, we allocate budget equally to each term:

$$\varepsilon_F,\ \varepsilon_R,\ \varepsilon_S \leq \varepsilon/3.$$

Substituting these into the bounds derived in Lemma 26 (gate complexity) and Lemma 25 (sample complexity), we obtain:

- **Gate complexity (per sample):**

$$C_{\text{gate}} = O\left(\Gamma \cdot \log\left(\frac{c(\varepsilon/3)}{\varepsilon}\right) \cdot \left(a_{\max}\Upsilon\lambda_{\text{comm}}\, T(\varepsilon/3)\right)^{1+\frac{1}{p}}\right),$$

  as detailed in Lemma 26.

- **Sample complexity:**

$$C_{\text{sample}} = O\left(\frac{c(\varepsilon/3)^4 \cdot (\log\log(1/\varepsilon))^4}{\varepsilon^2} \cdot \log\left(\frac{1}{\delta}\right)\right),$$

  as shown in Lemma 25.

Finally, the classical preprocessing cost is dominated by computing the Fourier coefficients $\{c_k\}$, Richardson coefficients $\{b_j\}$, and choosing approximation parameters for the product formula, which scales as $O(\log(1/\varepsilon))$. □

3.4.3 Estimating distributions

Our final task is to estimate a distribution corresponding to the state $f(H)\,|\psi\rangle$. Specifically, we wish to statistically approximate the vector $\sum_{\vec{z}_n\in(0,1)^n} \left|\langle\vec{z}_n|\, f(H)\,|\psi\rangle\right|^2 |\vec{z}_n\rangle$. We remark that this is proportional to the probability distribution that one obtains from sampling the state $f(H)\,|\psi\rangle\,/\|f(H)\,|\psi\rangle\,\|_2$ in the computational basis.

In Appendix A6 of [WMB24] the following Lemma is essentially given.

Lemma 28. *Suppose an operator G can be decomposed into a linear combination of unitaries as $G = \sum_i g_i U_i$. Using a generalized Hadamard test, we have an algorithm to return a vector $\vec{v}$ that, with probability at least $(1-\delta)$, satisfies*

$$\|\vec{v} - \vec{G}\|_2 \leq \varepsilon \tag{14}$$

using $O\left(\frac{g^4}{\varepsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ samples, where $g := \sum_i |g_i|$, and each generalized Hadamard test calls two unitaries from $\{U_i\}_i$.

The algorithm can be simply stated. First, we classically construct the probability distribution $p_i := |g_i|/g$. Then, sample two indices (i, j) from this distribution and implement the generalized Hadamard test in Figure 1 with U_i and U_j. Perform a measurement on all qubits in the computational basis, obtaining an n-qubit string which we denote $(z, \vec{z}_n) \in \{0,1\}^{n+1}$. Based on this outcome output

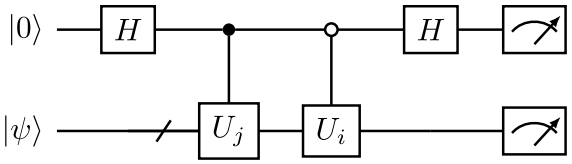

Figure 1: **Generalized Hadamard test.**

the vector $g^2\phi_{ij}(-1)^z\,|\vec{z}_n\rangle$ where we denote $\phi_{ij} = g_i g_j/|g_i g_j|$. Repeat the sampling procedure and take the sample mean of all outcomes.

Note that in the previous section we already found a operator G in this form that approximates $f(H)$ to additive error in operator norm, Thus, application of a triangle inequality allows the following theorem.

Theorem 29 (Distribution recovery)**.** *We give an algorithm for Task* (2) *with success probability at least* $1-\delta$ *and*

- ***Gate complexity (per sample):***

$$C_{\text{gate}} = O\left(\Gamma \cdot \log\left(\frac{c(\varepsilon/3)}{\varepsilon}\right) \cdot \left(a_{\max}\Upsilon\lambda_{\text{comm}}\,T(\varepsilon/3)\right)^{1+\frac{1}{p}}\right),$$

  where $T(\varepsilon/3) := \max_k |t_k|$ *is the largest time parameter in the Fourier approximation with error* $\varepsilon/3$*, and* $c(\varepsilon/3) := \|\vec{c}\|_1$ *is the associated* ℓ_1*-norm of the Fourier coefficients.*

- ***Sample complexity:***

$$C_{\text{sample}} = O\left(\frac{c(\varepsilon/3)^4 \cdot (\log\log(1/\varepsilon))^4}{\varepsilon^2} \cdot \log\left(\frac{1}{\delta}\right)\right),$$

- ***Classical preprocessing time:***

$$C_{\text{pre}} = O(K + \log(1/\varepsilon)).$$

4 k-local systems

In this section we discuss how to obtain commutator scaling for k-local systems for our algorithms, as well as other product formula extrapolation algorithms. A key technical tool is a based on an insight of Mizuta to exploit the fact that the BCH formula can be tightly truncated whilst maintaining good approximation of Trotterized time evolution [Miz26].

Lemma 30 (Truncated BCH formula [Miz26])**.** *For a* k*-local system on* n *qubits with maximum energy per site* g*, define the truncated effective Trotter Hamiltonian*

$$\widetilde{H}_{\text{eff}}(t,\varepsilon) := H + \sum_{j=1}^{p_0(\varepsilon)} E_{j+1} t^j\,. \tag{15}$$

If we set the truncation order as $p_0(\varepsilon) = \lceil \log(2n/\varepsilon) \rceil$ *for some parameter* $\varepsilon \in (0,1)$*, then we have*

$$\|e^{-iH_{\text{eff}}(t)t} - e^{-i\widetilde{H}_{\text{eff}}(t,\varepsilon)t}\| \leq \varepsilon\,. \tag{16}$$

The above lemma implies that for an extrapolation schedule $\{s_i\}_i$ with coefficients $\{b_i\}_i$, the approximation error can be bounded as

$$\Big\| \sum_i b_i e^{-iH_{\text{eff}}(s_i T)T} - \sum_i b_i e^{-i\widetilde{H}_{\text{eff}}(s_i T, \varepsilon_{\text{tr}}\hat{s}/\|\vec{b}\|_1)T} \Big\| \leq \varepsilon_{\text{tr}}\,, \tag{17}$$

where $\hat{s}$ is any given lower bound on all s_i, by using a triangle inequality and bounding the telescoping sum. That is, we can control the error by setting a truncation parameter $p_0(\varepsilon_{\text{tr}}\hat{s}/\|\vec{b}\|_1) = \lceil \log(2n\|\vec{b}\|_1/\hat{s}\varepsilon_{\text{tr}}) \rceil$. This will be useful to us as it means we only require knowledge of error operators to bounded order $p_0(\varepsilon_{\text{tr}}\hat{s}/\|\vec{b}\|_1)$, and allows us to avoid divergences in the extrapolated commutator factor. Specifically, for any error parameter ε we can simply repeat the same analysis of Lemma 10 for $\widetilde{H}_{\text{eff}}(t,\varepsilon)$ and obtain an error series

$$e^{-i\widetilde{H}_{\text{eff}}(sT,\varepsilon)} - e^{-iHT} = \sum_{j \in \sigma\mathbb{Z}_+ \geq p} s^j \tilde{E}_{j+1}(T,\varepsilon)\,, \tag{18}$$

with error operators now of the form

$$\tilde{E}_{j+1}(T,\varepsilon) \coloneqq \sum_{l=1}^{\lfloor j/p \rfloor} T^{j+l} \int_0^1 ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_{l-1}} ds_l \sum_{\substack{p \leq j_1 \ldots j_l \leq p_0(\varepsilon) \\ \in \sigma\mathbb{Z}_+ \\ j_1 + \cdots + j_l = j}} \left(\prod_{\kappa=l}^{1} e^{i(s_{\kappa-1} - s_\kappa)TH} i E_{j_\kappa + 1} \right) e^{i s_l T H}\,,$$

$$\|\tilde{E}_{j+1}(T,\varepsilon)\| \leq (a_{\max}\Upsilon T)^j \sum_{l=1}^{\lfloor j/p \rfloor} \frac{(a_{\max}\Upsilon T)^l}{l!} \sum_{\substack{p \leq j_1 \ldots j_l \leq p_0(\varepsilon) \\ \in \sigma\mathbb{Z}_+ \\ j_1 + \cdots + j_l = j}} \left(\prod_{\kappa=1}^{l} \frac{\alpha_{\text{comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2} \right) = \sum_{l=1}^{\lfloor j/p \rfloor} \frac{(a_{\max}\Upsilon T \lambda_{\text{comm}}^{(s\varepsilon_{\text{tr}})})^{j+l}}{l!}\,, \tag{19}$$

where $\lambda_{\text{comm}}^{(\varepsilon)}$ is defined as

$$\lambda_{\text{comm}}^{(\varepsilon)} := \sup_{\substack{j \in \sigma\mathbb{Z}_+ \geq \sigma m \\ 1 \leq l \leq \lfloor j/p \rfloor}} \left(\sum_{\substack{p \leq j_1 \ldots j_l \leq p_0(\varepsilon) \\ \in \sigma\mathbb{Z}_+ \\ j_1 + \cdots + j_l = j}} \left(\prod_{\kappa=1}^{l} \frac{\alpha_{\text{comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2} \right) \right)^{\frac{1}{j+l}}\,. \tag{20}$$

Lemma 31. *For a* k*-local Hamiltonian* H *on* n *qubits with maximum on site energy* g*, we have*

$$\lambda_{\text{comm}}^{(\varepsilon)} = O\Big(kg(p\Lambda^{1/p} + \log(n/\varepsilon)) \Big)\,. \tag{21}$$

Proof. The analysis is identical to that found in [Miz26, Lemma 5], and we provide a proof here for completeness. For k-local systems, we have $\alpha_{\text{comm}}^{(j)} \leq (j-1)!(2kg)^{j-1}\Lambda$ [Chi+21b]. Thus, we can write

$$\sum_{\substack{p\leq j_1\ldots j_l\leq p_0(\varepsilon)\\ \in\sigma\mathbb{Z}_+\\ j_1+\cdots+j_l=j}}\left(\prod_{\kappa=1}^{l}\frac{\alpha_{\text{comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2}\right)\leq \sum_{\substack{p\leq j_1\ldots j_l\leq p_0(\varepsilon)\\ \in\sigma\mathbb{Z}_+\\ j_1+\cdots+j_l=j}}\left(\prod_{\kappa=1}^{l}(j_\kappa-2)!(2kg)^{j_\kappa}\Lambda/g\right) \tag{22}$$

$$\leq \sum_{\substack{p\leq j_1\ldots j_l\leq p_0(\varepsilon)\\ \in\sigma\mathbb{Z}_+\\ j_1+\cdots+j_l=j}}\prod_{\kappa=1}^{l}\left(2kgj_\kappa(\Lambda/g)^{1/j_\kappa}\right)^{j_\kappa} \tag{23}$$

$$\leq \left(\max_{p\leq j'\leq p_0(\varepsilon)}\left(2kgj'(\Lambda/g)^{1/j'}\right)\right)^{j+l}\sum_{\substack{j_1\ldots j_l\geq 0\\ j_1+\cdots+j_l=j}}1 \tag{24}$$

$$\leq (4kg)^{j+l}\left(\max_{p\leq j'\leq p_0(\varepsilon)}\left(j'(\Lambda/g)^{1/j'}\right)\right)^{j+l} \tag{25}$$

where in the final line we have used the stars-and-bars formula to bound the sum by $\binom{j+l-1}{l-1}\leq 2^{j+l}$. The function $x(\Lambda/g)^{1/x}$ monotonically decreases in x for all $0<x<\log(\Lambda/g)$ and increasing for $x\geq\log(\Lambda/g)$. Thus, we have

$$\max_{p\leq j'\leq p_0(\varepsilon)}\left(j'(\Lambda/g)^{1/j'}\right)\leq \max\left\{p\Lambda^{1/p}, p_0(\varepsilon)\Lambda^{1/p_0(\varepsilon)}\right\}=O\left(p\Lambda^{1/p}+p_0(\varepsilon)\right), \tag{26}$$

due to the fact that $\Lambda\leq ng$ and $x^{1/\log x}=O(1)$. Putting everything together, we have

$$\lambda_{\text{comm}}^{(\varepsilon)}\leq 4kg(p\Lambda^{1/p}+p_0(\varepsilon)), \tag{27}$$

as required. □

Lemma 32 (Generic Richardson extrapolation error for time signal, k- local)**.** *Let $\mathcal{P}$ be a staged pth order product formula of symmetry class σ, where $\sigma=2$ if $\mathcal{P}$ is symmetric, 1 otherwise. For a target evolution time T let*

$$\mathcal{P}_{p,m}^{(R)}(T):=\sum_{k=1}^{m}b_i\mathcal{P}^{1/s_i}(s_iT)=\sum_{i=1}^{m}b_ie^{-iH_{\text{eff}}(s_iT)T} \tag{28}$$

denote an m-term Richardson extrapolation, with ascending sequence of Trotter steps $r_k=1/s_k\in\mathbb{Z}_+$, which cancel the powers $s^\sigma, s^{2\sigma},\ldots,s^{\sigma(m-1)}$. We consider a Hamiltonian which assumes the conditions of Lemma 10. Then, the error in the extrapolation of a Trotterized time signal, as compared to an exact time signal, satisfies

$$\left|\operatorname{Tr}[\rho(\mathcal{P}_{p,m}^{(R)}(T)-e^{-iHT})]\right|\leq 4\|\vec{b}\|_1\eta^{\lfloor\sigma m/p\rfloor}\left(s_1a_{\max}\Upsilon\lambda_{\text{comm}}^{(\hat{s}\varepsilon/2\|\vec{b}\|_1)}T\right)^{\sigma m}+\varepsilon/2\,.$$

for any quantum state ρ, where $\|\vec{b}\|_1=\sum_k|b_k|$, and $\hat{s}$ is a lower bound on all s_k.

Proof. Set the truncation parameter as $p_0(\varepsilon') = p_0(\varepsilon\hat{s}/2\|\vec{b}\|_1) = \lceil \log(4n\|\vec{b}\|_1/\hat{s}\varepsilon)\rceil$. The extrapolation error takes the form

$$\mathrm{Tr}[\rho(\mathcal{P}^{(R)}_{p,m}(T) - e^{-iHT})] = \sum_{k=1}^{m} b_k \|R_{\sigma(m-1)}(T, s_k, \varepsilon')\|. \tag{29}$$

where $R_{\sigma(m-1)}(T, s_k)$ denotes the Taylor remainder of degree $\sigma(m-1)$ which takes the form

$$R_{\sigma(m-1)}(T, s_k) := \sum_{\substack{j\in\sigma\mathbb{Z}_+ \\ j\geq\sigma m}} s_k^j \tilde{E}_{j+1}(T, \varepsilon')\,, \tag{30}$$

for each inverse Trotter step s_k. The size of this is bounded as

$$\begin{aligned}\left\|R_{\sigma(m-1)}(T, s_k, \varepsilon')\right\| &\leq \sum_{\substack{j\in\sigma\mathbb{Z}_+ \\ j\geq\sigma m}} s_k^j \left\|\tilde{E}_{j+1}(T, \varepsilon')\right\| \\ &\leq \sum_{\substack{j\in\sigma\mathbb{Z}_+ \\ j\geq\sigma m}} s_k^j \sum_{l=1}^{\lfloor j/p\rfloor} \frac{\left(a_{\max}\Upsilon\lambda^{(\varepsilon')}_{\mathrm{comm}}T\right)^{j+l}}{l!}\,,\end{aligned}$$

where as before we have used Lemma 10 and denoted $\lambda_{j,l}$ as in the statement of this Lemma. As $\lambda^{(\varepsilon')}_{\mathrm{comm}}$ is independent of the extrapolation index i, we can proceed as before and bound the extrapolation error simply as

$$\left|\mathrm{Tr}[\rho(\mathcal{P}^{(R)}_{p,m}(T) - e^{-iHT})]\right| \leq \|\vec{b}\|_1 \cdot \left\|R_{\sigma(m-1)}(T, s_1)\right\| \leq 4\|\vec{b}\|_1\eta^{\lfloor\sigma m/p\rfloor}\left(s_1 a_{\max}\Upsilon\lambda_{\mathrm{comm}}T\right)^{\sigma m}\,.$$

on the condition that $s_1(a_{\max}\Upsilon\lambda_{\mathrm{comm}}T)^{1+1/p} \leq 1/2$. Finally, one can check that setting $\varepsilon' = \varepsilon\hat{s}/2\|\vec{b}\|_1$ guarantees that the truncation error is bounded as

$$\left\|\mathcal{P}^{(R)}_{p,m}(T) - \sum_i b_i e^{-i\widetilde{H}_{\mathrm{eff}}(s_i T, \varepsilon\hat{s}/2\|\vec{b}\|_1)T}\right\| \leq \varepsilon/2\,, \tag{31}$$

and the statement of the lemma follows by a triangle inequality. □

The above lemma allows us to repeat prior analysis with $\lambda^{(\hat{s}\varepsilon/2\|\vec{b}\|_1)}_{\mathrm{comm}}$ for k-local systems in place of λ_{comm}. In the case of compiling time evolution, we have $\hat{s} = \mathrm{poly}(n, g, T, \varepsilon^{-1})$ (note that $\Lambda \leq ng$).

Moreover, we have $\lambda^{(\hat{s}\varepsilon/2\|\vec{b}\|_1)}_{\mathrm{comm}} \leq \widetilde{\lambda}_{k\text{-local}}$, where

$$\widetilde{\lambda}_{k\text{-local}} = O\Big(kg(p\Lambda^{1/p} + g\log(ngT/\varepsilon))\Big)\,.$$

Further, when compiling other functions T is simply taken to be the maximum time parameter of the Fourier series, and an additive term of $g\log\log(c(\varepsilon/3))$ is inherited for the Fourier weight $c(\varepsilon/3)$.

5 Fermionic systems with fixed particle number

In many quantum simulation settings (such as electronic structure problems and lattice fermion models) the dynamics are confined to fixed particle-number subspaces. That is, the Hamiltonians and observables of interest are number-conserving, mapping η-fermion states to other η-fermion states. This constraint enables a more refined error analysis when applying Algorithm 1 if the state in question is of well-defined fermion number. Specifically, we introduce the fermionic η-seminorm, which is effectively the operator norms projected to an η-particle sector. This seminorm often yields significantly tighter bounds for Trotter and extrapolation errors, as the relevant nested commutators typically have much smaller fermionic seminorms than their full operator norms. As a result, the gate complexity of simulation algorithms can be reduced by conducting the error analysis in terms of fermionic seminorms, using improved constants $\alpha_{\text{comm}}^{(\eta)} \leq \alpha_{\text{comm}}$ established in recent works [MCS22; SHC21; Low+23].

5.1 Closure under Fermionic Seminorm

We work in a fermionic Fock space with fixed particle number sectors labeled by η. For any operator X that preserves particle number, we define the fermionic η-seminorm:

$$\|X\|_\eta := \max_{|\psi_\eta\rangle, |\phi_\eta\rangle} \frac{|\langle\phi_\eta|X|\psi_\eta\rangle|}{\||\psi_\eta\rangle\| \cdot \||\phi_\eta\rangle\|},$$

where $|\psi_\eta\rangle$ and $|\phi_\eta\rangle$ are η-electron states. This seminorm quantifies the action of X restricted to the η-particle subspace.

Lemma 33 (Fermionic seminorm as a projected spectral norm)**.** *For any number-preserving operator X, the fermionic η-seminorm satisfies*

$$\|X\|_\eta = \max_{|\psi_\eta\rangle, |\phi_\eta\rangle} |\langle\phi_\eta|X|\psi_\eta\rangle| = \|X\Pi_\eta\|,$$

where Π_η is the projector onto the η-electron subspace.

We now collect closure properties of number-preserving operators, which will be critical in bounding errors using the η-seminorm.

Lemma 34 (Closure under products)**.** *If A and B are number-preserving, then AB is also number-preserving.*

Lemma 35 (Closure under linear combinations)**.** *If $\{X_m\}$ are number-preserving and $c_m \in \mathbb{C}$, then $\sum_m c_m X_m$ is number-preserving.*

Lemma 36 (Exponentials of number-preserving Hamiltonians)**.** *Let*

$$\hat{N} := \sum_{\upsilon,\gamma} \hat{a}^\dagger_{(\upsilon,\gamma)} \hat{a}_{(\upsilon,\gamma)}$$

be the total number operator. If $H = \sum_\ell H_\ell$ is a Hamiltonian with each H_ℓ number-preserving, then for any real θ, the unitary $e^{i\theta H}$ is number-preserving.

Lemma 37 (Commutators preserve particle number)**.** *If A and B are number-preserving, then so is $[A, B] := AB - BA$. More generally, any nested commutator of number-preserving operators is number-preserving.*

Lemma 38 (Error terms preserve particle number)**.** *Let $\tilde{E}_{j+1,K}(T)$ be defined via finite products and nested commutators of number-preserving local Hamiltonian terms $\{H_\ell\}$. Then each $\tilde{E}_{j+1,K}(T)$ is number-preserving.*

Proof. By Lemmas 34, 35, and 37, number-preserving operators are closed under products, linear combinations, and commutators. Since each $\tilde{E}_{j+1,K}(T)$ is constructed from such operations on $\{H_\ell\}$, it must also be number-preserving. □

5.2 Bounding with Fermionic Seminorm

Having established the essential closure properties of the fermionic seminorm, we now leverage this structure to derive tighter bounds on the coefficients in our error series expansion—and, ultimately, on the gate complexity.

Lemma 39 (Fermionic Error Formula with η-Seminorms)**.** *Let $\mathcal{P}$ be a staged pth order product formula. Assume all operators, including time-evolution operators, commutators, and effective Hamiltonians, are number-preserving. For any $s \in \mathbb{R}$, define*

$$\mathcal{P}(t) = e^{-itH_{\text{eff}}(t)}$$

as the approximate evolution operator over total time $T \in \mathbb{R}$, with Trotter step size $t = sT$. The case $s = 0$ is defined via the limit.

Suppose there exists $J \in \mathbb{Z}_+$ and $C \in \mathbb{R}_+$ such that

$$\sup_{j \geq J} \alpha^{(j)}_{\text{comm}} \left(a_{\max} \Upsilon |sT|\right)^j \leq C,$$

where $a_{\max} \coloneqq \max_{\upsilon,\gamma} |a_{(\upsilon,\gamma)}|$, and Υ bounds locality or support size.

Define $\sigma = 2$ if $\mathcal{P}$ is symmetric, and $\sigma = 1$ otherwise. Then for any $K \in \mathbb{Z}_+$, the approximation error in $\mathcal{P}^{1/s}(sT)$ compared to the exact evolution satisfies

$$\operatorname{tr}\left[\rho \left(\mathcal{P}^{1/s}(sT) - e^{-iHT}\right)\right] = \sum_{j \in \sigma\mathbb{Z}_+ \geq p} s^j \operatorname{tr}\left[\rho\, \tilde{E}_{j+1,K}(T)\right] + \operatorname{tr}\left[\rho\, \tilde{F}_K(T,s)\right],$$

where the fermionic seminorms satisfy the bounds

$$\|\tilde{E}_{j+1,K}(T)\|_\eta \leq \left(a_{\max}\Upsilon T\right)^j \sum_{l=1}^{\min\{K-1, \lfloor j/p \rfloor\}} \frac{\left(a_{\max}\Upsilon T\right)^l}{l!} \sum_{\substack{j_1,\ldots,j_l \in \sigma\mathbb{Z}_+ \geq p \\ j_1+\cdots+j_l = j}} \prod_{\kappa=1}^{l} \frac{\alpha^{(j_\kappa+1)}_{\eta\text{-comm}}}{(j_\kappa+1)^2},$$

$$\|\tilde{F}_K(T,s)\|_\eta \leq \frac{\left(a_{\max}\Upsilon T\right)^K}{K!} \sum_{j \in \sigma\mathbb{Z}_+ \geq Kp} \left(a_{\max}\Upsilon sT\right)^j \sum_{\substack{j_1,\ldots,j_K \in \sigma\mathbb{Z}_+ \geq p \\ j_1+\cdots+j_K = j}} \prod_{\kappa=1}^{K} \frac{\alpha^{(j_\kappa+1)}_{\eta\text{-comm}}}{(j_\kappa+1)^2}.$$

Define the scale parameter

$$\lambda_{\eta,j,l} := \left(\sum_{\substack{j_1,\ldots,j_l \in \sigma\mathbb{Z}_+ \geq p \\ j_1+\cdots+j_l=j}} \prod_{\kappa=1}^{l} \frac{\alpha_{\eta\text{-comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2} \right)^{\frac{1}{j+l}}.$$

Proof. Recall from [WW24b, Lemma 3] that a pth-order staged product formula has an error operator

$$\Delta(s) = \sum_{j \in \sigma\mathbb{Z}_+ \geq p} E_{j+1} \cdot (sT)^j,$$

where all E_j are number-preserving operators, and $\sigma = 2$ if $\mathcal{P}$ is symmetric, 1 otherwise.

Using the variation-of-parameters formula, we recursively expand the approximate time evolution operator $e^{iTH_{\text{eff}}(s)}$ as:

$$\begin{aligned} e^{iTH_{\text{eff}}(s)} = e^{iHT} &\\ &+ \sum_{l=1}^{K-1} \int_0^T d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{l-1}} d\tau_l \, e^{i(T-\tau_1)H} (i\Delta(s)) \cdots (i\Delta(s)) e^{iH\tau_l} \\ &+ \int_0^T \cdots \int_0^{\tau_{K-1}} d\tau_K \, e^{i(T-\tau_1)H} (i\Delta(s)) \cdots (i\Delta(s)) e^{i\tau_K H_{\text{eff}}(s)}. \end{aligned}$$

By Lemma 38, all time-evolution and commutator terms preserve fermionic number, so the entire expansion preserves it.

Expanding each $\Delta(s)$ into its power series,

$$\sum_{l=1}^{K-1} \int_0^T d\tau_1 \cdots \int_0^{\tau_{l-1}} d\tau_l \sum_{j_1,\ldots,j_l \in \sigma\mathbb{Z}_+ \geq p} (sT)^{j_1+\cdots+j_l} \prod_{\kappa=1}^{l} e^{i(\tau_{\kappa-1}-\tau_\kappa)H} iE_{j_\kappa+1} e^{iH\tau_l} = \sum_{j \in \sigma\mathbb{Z}_+ \geq p} s^j \tilde{E}_{j+1,K}(T),$$

where $\tau_0 := T$ and

$$\tilde{E}_{j+1,K}(T) := \sum_{l=1}^{\min\{K-1, \lfloor j/p \rfloor\}} T^{j+l} \int_0^1 ds_1 \cdots \int_0^{s_{l-1}} ds_l \sum_{\substack{j_1+\cdots+j_l=j \\ j_i \in \sigma\mathbb{Z}_+ \geq p}} \prod_{\kappa=1}^{l} e^{i(s_{\kappa-1}-s_\kappa)TH} iE_{j_\kappa+1} e^{is_l TH}.$$

Using the unitary invariance and submultiplicativity of the fermionic seminorm, we get

$$\begin{aligned} \|\tilde{E}_{j+1,K}(T)\|_\eta &\leq \sum_{l=1}^{\min\{K-1, \lfloor j/p \rfloor\}} T^{j+l} \int_0^1 ds_1 \cdots \int_0^{s_{l-1}} ds_l \sum_{j_1+\cdots+j_l=j} \prod_{\kappa=1}^{l} \|E_{j_\kappa+1}\|_\eta \\ &\leq T^j \sum_{l=1}^{\min\{K-1, \lfloor j/p \rfloor\}} \frac{T^l}{l!} \sum_{j_1+\cdots+j_l=j} \prod_{\kappa=1}^{l} \|E_{j_\kappa+1}\|_\eta. \end{aligned}$$

Using the effective Hamiltonian error expansion bound

$$\|E_{j+1}\|_\eta \leq \alpha_{\text{comm}}^{(j+1)} \frac{(a_{\max}\Upsilon)^{j+1}}{(j+1)^2},$$

we conclude

$$\|\tilde{E}_{j+1,K}(T)\|_\eta \leq (a_{\max}\Upsilon T)^j \sum_{l=1}^{\min\{K-1,\lfloor j/p \rfloor\}} \frac{(a_{\max}\Upsilon T)^l}{l!} \sum_{j_1+\cdots+j_l=j} \prod_{\kappa=1}^{l} \frac{\alpha_{\eta\text{-comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2}.$$

For the remainder term $\tilde{F}_K(T,s)$, applying the triangle inequality and unitarity,

$$\begin{aligned}
\|\tilde{F}_K(T,s)\|_\eta &\leq \frac{T^K}{K!}\|\Delta(s)\|_\eta^K \leq \frac{T^K}{K!}\left(\sum_{j\in\sigma\mathbb{Z}_+\geq p} \|E_{j+1}\|_\eta (sT)^j\right)^K \\
&= \frac{T^K}{K!} \sum_{j\in\sigma\mathbb{Z}_+\geq Kp} (sT)^j \sum_{j_1+\cdots+j_K=j} \prod_{\kappa=1}^{K} \|E_{j_\kappa+1}\|_\eta \\
&\leq \frac{(a_{\max}\Upsilon T)^K}{K!} \sum_{j\in\sigma\mathbb{Z}_+\geq Kp} (a_{\max}\Upsilon sT)^j \sum_{j_1+\cdots+j_K=j} \prod_{\kappa=1}^{K} \frac{\alpha_{\eta\text{-comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2}.
\end{aligned}$$

□

Theorem 40. *For fermionic systems, choose the largest inverse Trotter step number s_1 such that such that $s_1 a_{\max}\Upsilon\lambda_{\eta comm}T < 1/2$. Define $\eta := \max\{1, a_{max}\Upsilon T\lambda_{\eta comm}\}$. The m-term Richardson extrapolation error for a staged pth-order product formula of symmetry class σ is bounded as*

$$\left|Tr[\rho(\mathcal{P}_{p,m}^{(R)}(T) - e^{-iHT})]\right| \leq 4\|\vec{b}\|_1 \eta^{\lceil \sigma m/p\rceil} \left(s_1 a_{\max}\Upsilon\lambda_{\eta\text{-}comm}T\right)^{\sigma m}$$

Proof. Assume that our state ρ exists in the η electron space, therefore $\rho = \rho\prod_\eta$. Note that we can set

$$\begin{aligned}
\text{Tr}[\rho\Pi_\eta\tilde{E}_{j+1,K}(T)] &\leq \|\Pi_\eta\tilde{E}_{j+1,K}(T)\|_\infty\|\rho\|_1 = \|\Pi_\eta\tilde{E}_{j+1,K}(T)\|_\infty = \|\tilde{E}_{j+1,K}(T)\|_\eta \\
\text{Tr}[\rho\Pi_\eta\tilde{F}_K(T,s)] &\leq \|\Pi_\eta\tilde{F}_K(T)\|_\infty\|\rho\|_1 = \|\Pi_\eta\tilde{F}_K(T,s)\|_\infty = \|\tilde{F}_K(T,s)\|_\eta
\end{aligned}$$

where the inequality is due to the Holder's trace matrix inequality. Since we can bound

$$\begin{aligned}
\text{Tr}[\rho\Pi_\eta\tilde{E}_{j+1,K}(T)] &\leq \|\tilde{E}_{j+1,K}(T)\|_\eta \\
\text{Tr}[\rho\Pi_\eta\tilde{F}_K(T,s)] &\leq \|\tilde{F}_K(T,s)\|_\eta
\end{aligned}$$

we can use the bounds provided in Lemma 39 and the same techniques from Lemma 12 to obtain

$$\left|\text{Tr}[\rho(\mathcal{P}_{p,m}^{(R)}(T) - e^{-iHT})]\right| \leq 4\|\vec{b}\|_1 \eta^{\lceil \sigma m/p\rceil} \left(s_1 a_{\max}\Upsilon\lambda_{\eta\text{-comm}}T\right)^{\sigma m}$$

□

Given this final form, we can now substitute the recently derived bounds on $\alpha_{\eta\text{-comm}}$ to obtain tighter guarantees than those available for more general systems. In particular, the work of [Low+23] provides the following sharp error bounds for Trotterization in the context of fermionic Hamiltonians:

Lemma 41 (Trotter error with fermionic induced 1-norm scaling ([Low+23], Theorem 4))**.** *Let $H = T + V := \sum_{j,k} \tau_{j,k} A_j^\dagger A_k + \sum_{l,m} v_{l,m} N_l N_m$ be an interacting-electronic Hamiltonian, and let $\mathcal{P}$ denote a pth-order product formula that splits the evolution under T and V. Then,*

$$\left\| \mathcal{P}(t) - e^{-itH} \right\|_\eta = O\left(\left(|||\tau|||_1 + |||v|||_{1,[\eta]} \right)^{p-1} |||\tau|||_1 |||v|||_{1,[\eta]} \, \eta \, t^{p+1} \right).$$

Here, the (restricted) fermionic induced 1-norms are defined as

$$|||\tau|||_1 = \max_j \sum_k |\tau_{j,k}|, \quad |||v|||_{1,[\eta]} = \max_j \max_{k_1 < \cdots < k_\eta} \left(|v_{j,k_1}| + \cdots + |v_{j,k_\eta}| \right).$$

6 Partially Randomized product formulae

We now apply our Richardson extrapolation framework to the partially randomized product formulae introduced in [Gün+25], which combine the structure of deterministic Trotter formulas with randomized simulation techniques.

6.1 Motivation and Prior Work

Randomized techniques have proven useful in Hamiltonian simulation, especially for mitigating worst-case behavior in certain settings [WBC22b]. However, for a Hamiltonian $H = \sum_{i=1}^{\Gamma} H_i$, fully randomized methods tend to scale poorly with the inverse error ε^{-1} and the total norm $\lambda = \sum_{i=1}^{\Gamma} \|H_i\|$. On the other hand, deterministic product formulae often scale poorly with Γ, the number of terms.

To leverage the advantages of both approaches, we consider a hybrid scheme in which part of the Hamiltonian is simulated deterministically and the remainder is handled using randomization. Specifically, we adopt the decomposition proposed in [Gün+25]:

$$H = \underbrace{\sum_{l=1}^{L_A} H_l}_{H_A} + \underbrace{\sum_{m=1}^{M} h_m P_m}_{H_B}, \tag{32}$$

where H_A is treated deterministically and H_B is simulated via randomized techniques. Each P_m is assumed to be a Hermitian unitary (e.g., a Pauli operator), so that $P_m^2 = I$.

This decomposition is designed to satisfy:

$$\lambda_B := \sum_{m=1}^{M} |h_m| \ll \lambda, \quad \text{and} \quad L_A \ll M,$$

where $\lambda = \sum_{l=1}^{L_A} \|H_l\| + \sum_{m=1}^{M} |h_m|$ is the total 1-norm of H. Intuitively, this ensures that most of the Hamiltonian's norm resides in the smaller, randomized component H_B, while keeping the deterministic portion H_A lightweight.

6.2 Gate Complexity Improvement

The partially randomized framework leads to a favorable gate complexity bound when combined with higher-order product formulae. The key result from [Gün+25] is summarized below.

Lemma 42 (Gate Complexity of Partially Randomized product formulae [Gün+25, Theorem V.1])**.** *Let $H = H_A + H_B$ be decomposed as in (32), and suppose a pth-order product formula is applied over r steps with step size $t = T/r$, where*

$$r = O\left((\alpha_{\text{comm}}^{(p+1)})^{1/p} T^{1+1/p} \varepsilon^{-1/p}\right).$$

Then the total gate complexity is

$$O\left(rL_A + \lambda_B^2 T^2\right).$$

The first term corresponds to the deterministic simulation of H_A using product formulae, while the second term arises from approximating the randomized evolution $e^{iH_B t}$ via a truncated Taylor series.

6.3 LCU Decomposition of Randomized Taylor Expansion

Before we introduce how we use Richardson-extrapolation for our partially random circuits, we share a useful lemma that allows us to estimate time evolution operator as a linear combination of random unitaries. This result of this lemma is used in Algorithm 2. The is the core lemma that the partial randomization approach exploits.

Lemma 43 (Approximate LCU decomposition via Randomized Taylor Expansion of Time Evolution ([WBC22a], Lemma 2))**.** *Let $H = \sum_k h_k P_k$ be a Hermitian operator, where P_k are Pauli operators and h_k are real coefficients. Define $\lambda := \sum_k |h_k|$. Then, for any time $T > 0$ and any integer $d \geq 1$, the truncated Taylor expansion of e^{-iHT} to order d can be written as a Linear Combination of Unitaries (LCU) of the form:*

$$\sum_{r=0}^{d} \frac{(iHT)^r}{r!} = \beta \sum_j q_j U_j,$$

where:

- *$\{q_j\}_j$ is a probability distribution: $q_j \geq 0$ and $\sum_j q_j = 1$.*
- *Each U_j is a unitary consisting of a product of at most d Pauli operators (i.e., $U_j = P_{k_1} \cdots P_{k_r}$ with $r \leq d$).*
- *The normalization factor β satisfies $\beta \leq \exp\left(\frac{\lambda^2 T^2}{d}\right)$.*

We now describe how to apply Richardson extrapolation to partially randomized product formulae, and show that (similar to Algorithm 1) this improves the error scaling from $\varepsilon^{-1/p}$ to $\log(1/\varepsilon)$.

To maintain notational consistency throughout, we define the partially randomized product formula $\mathcal{S}_p^{1/s}(t)$ (as opposed to $\mathcal{P}$ from before) corresponding to a decomposition of the Hamiltonian H into $\{H_l\}_{l=1}^{L_A} \cup \{H_B\}$, where the total simulation time is $t = sT$ and T is the target evolution time. In this notation, we write:

$$\mathcal{S}_p^{1/s}(t) = \left(\prod_{l=1}^{L_A} e^{iH_l t} \right) e^{iH_B t}.$$

We also define the m-term Richardson extrapolated approximation of the full evolution as:

$$\mathcal{S}_{p,m}^{(R)}(T) := \sum_{k=1}^{m} b_k \, \mathcal{S}^{1/s_k}(s_k T),$$

Algorithm 2 Richardson-Extrapolated Partially Randomized Product Formulae

Input:

- Hamiltonian $H = \sum_{l=1}^{L_A} H_l + \sum_{m=1}^{M} h_m P_m$
- Richardson step sizes $\{s_k\}_{k=1}^{m}$
- Richardson coefficients $\{b_k\}_{k=1}^{m}$ such that: $\mathrm{Tr}\left[Z\mathcal{S}_{p,m}^{(R)}(T)\right] = \sum_{k=1}^{m} b_k \cdot \mathrm{Tr}\left[Z\mathcal{S}(s_k T)\right]$
- Total simulation time T

Output: Approximation $\hat{Y}_N$ to $\mathrm{Tr}[Ze^{-iHT}]$ within error ε

Initialize $Y_{\mathrm{sum}} \leftarrow 0$
for $i = 1$ to N **do**
  Sample k_i from the distribution $\left\{\frac{|b_k|}{b}\right\}$ where $b = \sum_{k=1}^{m} |b_k|$
  Set $r \leftarrow \frac{1}{s_{k_i}}$, $\delta \leftarrow \frac{T}{r}$, and construct $\mathcal{P}(\delta)$ using H_A
  Set $t \leftarrow s_{k_i} T$, $d \leftarrow \frac{\lambda_B^2 t^2}{r}$, and $\tau \leftarrow \frac{t}{d}$
  for $j = 1$ to r **do**
    Use randomized Taylor expansion from Lemma 43: $e^{iH_B t} \approx \beta(\tau)^d \sum_m \tilde{q}_m W_m$
    Sample W_m from the distribution $\left\{\frac{|q_m|}{q}\right\}$
    Construct a Trotter step using $\mathcal{P}(\delta)$ and sampled W_m
  end for
  Let $\mathcal{S}_{k_i}$ be the full r-step circuit built from $\mathcal{P}(\delta)$ and W_m
  Run Hadamard test circuits to estimate:
  - $\mathrm{Re}\left[\mathrm{Tr}(Z\mathcal{S}(s_{k_i}T))\right] \to X_R^{(i)} \in [-1, 1]$
  - $\mathrm{Im}\left[\mathrm{Tr}(Z\mathcal{S}(s_{k_i}T))\right] \to X_I^{(i)} \in [-1, 1]$
  Let $X^{(i)} \leftarrow X_R^{(i)} + i \cdot X_I^{(i)}$
  Set $Y^{(i)} \leftarrow X^{(i)} \cdot \mathrm{sign}(b_{k_i}) \cdot b$
  Update $Y_{\mathrm{sum}} \leftarrow Y_{\mathrm{sum}} + Y^{(i)}$
end for
Return: $\hat{Y}_N = \frac{1}{N} Y_{\mathrm{sum}}$

Theorem 44 (Extrapolated Partial Randomization)**.** *Consider a matrix decomposed as $H = H_A + H_B$ as in Eq. (32). Let $\tilde{\lambda}_{comm}$ be the commutator bound corresponding to the decomposition of H into $\{H_l\}_{l=1}^{L_A} \cup H_B$ and let Z be an operator with bounded Schatten-1 norm $\|Z\|_1 \leq 1$ such that $\mathrm{Tr}[ZU]$ is implementable via a Hadamard test given (controlled) U. Then, using Algorithm 2 we can obtain an ε-additive approximation to $\mathrm{Tr}[Ze^{-iHT}]$ with the following resource complexities:*

- ***Circuit samples:***

$$C_{sample} = O\left(\frac{(\log\log(1/\varepsilon))^2}{\varepsilon^2} \log\left(\frac{1}{\delta}\right)\right),$$

- ***Gates per circuit:***

$$C_{gate} = O\left(L_A(a_{\max}\Upsilon\tilde{\lambda}_{\text{comm}}T)^{1+\frac{1}{p}}\log(1/\varepsilon) + \lambda_B^2 T^2\right).$$

Proof. Let $\hat{Y}_N$ denote the empirical estimate returned by Algorithm 2 after N samples. The expectation of this estimate is

$$\mathbb{E}[\hat{Y}_N] = \text{Tr}\left[Z\,\mathcal{S}_{p,m}^{(R)}(T)\right].$$

We decompose the total error into sampling ε_S and Richardson error ε_R:

$$\left|\text{Tr}[Ze^{-iHT}] - \hat{Y}_N\right| \leq \underbrace{\left|\hat{Y}_N - \text{Tr}\left[Z\,\mathcal{S}_{p,m}^{(R)}(T)\right]\right|}_{\varepsilon_{\text{S}}} + \underbrace{\left|\text{Tr}\left[Z\,\mathcal{S}_{p,m}^{(R)}(T)\right] - \text{Tr}\left[Ze^{-iHT}\right]\right|}_{\varepsilon_{\text{R}}}.$$

To guarantee overall error at most ε, we set: $\varepsilon_{\text{S}} \leq \frac{\varepsilon}{2}, \quad \varepsilon_{\text{R}} \leq \frac{\varepsilon}{2}$.
Sample Complexity. We estimate $\text{Tr}[Ze^{-iHT}]$ using an m-term Richardson extrapolation:

$$\mathcal{S}_{p,m}^{(R)}(T) = \sum_{k=1}^{m} b_k\,\mathcal{S}(s_k T),$$

where each $\mathcal{S}(s_k T)$ is a product formula for the Hamiltonian decomposed as $H_1, \ldots H_{L_A}, H_B$ and b_k are extrapolation coefficients. With each iteration of the loop $i = 1 \ldots N$, Algorithm 2 samples $X^{(i)}, Y^{(i)}$ where:

- $X^{(i)}$ is the sum of two Hadamard tests corresponding to the unitary circuit $\mathcal{S}(s_{k_i}T)$. Therefore mathematically we get that $\mathbb{E}[X^{(i)}] = \text{Tr}[Z\mathcal{S}(s_{k_i}T)]$
- $Y^{(i)} = \text{sign}(b_k)\cdot b\cdot X^{(i)}$.
- Therefore for the final the outputted $\hat{Y}_N$, we get that

$$\mathbb{E}[\hat{Y}_N] = \sum_{k=1}^{m}\frac{|b_k|}{b}\mathbb{E}[Y^{(i)}] = \sum_{k=1}^{m}\frac{|b_k|}{b}\mathbb{E}[\text{sign}(b_k)\cdot b\cdot X^{(i)}] = \sum_{k=1}^{m} b_k\text{Tr}[Z\mathcal{S}(s_{k_i}T)] = \text{Tr}[Z\mathcal{S}_{p,m}^{(R)}(T)]$$

By Hoeffding's inequality, to ensure that

$$\mathbb{P}(|\hat{Y}_N - \mathbb{E}[\hat{Y}_N]| \geq \varepsilon/2) \leq \delta,$$

Now consider $|Y^{(i_k)}| \leq M := b \leq O(\log\log(1/\varepsilon))$. it suffices that:

$$N \geq \frac{2M^2}{\varepsilon^2}\log\left(\frac{2}{\delta}\right).$$

We now bound M. By Lemma 13, the sum of extrapolation weights satisfies $b := \sum_k |b_k| = O(\log m) = O(\log\log(1/\varepsilon))$. Thus:

$$M \leq \|Z\| \cdot b = O(\log m) = O(\log\log(1/\varepsilon)).$$

So the sample complexity is:

$$C_{\text{sample}} = O\left(\frac{(\log\log(1/\varepsilon))^2}{\varepsilon^2} \log\left(\frac{1}{\delta}\right)\right).$$

Extrapolation error. By Lemma 16, the total observable error from the staged partially randomized Trotter formula satisfies:

$$\left|\text{Tr}[Ze^{-iHT}] - \text{Tr}[Z\mathcal{S}^{(R)}_{p,m}(T)]\right| \leq \frac{\varepsilon}{2},$$

provided that each stage uses at most $r_k \leq r_{\max}$ Trotter steps, where:

$$r_{\max} = O\left((a_{\max}\Upsilon\lambda_{\text{comm}}T)^{1+\frac{1}{p}} \log^2(1/\varepsilon)\right).$$

Gate complexity. Lemma 42 says that for a given Trotter circuit $\mathcal{S}(s_{k_i}T)$ obtained in step i of Algorithm 2 composed of r_{k_i} trotter steps, we get a gate complexity of $O(r_{k_i}L_A + \lambda_B^2T^2)$. Thus, the worst-case gate complexity of the algorithm (per quantum circuit) is:

$$C_{\text{gate}} = O\left(r_{\max}L_A + \lambda_B^2T^2\right) = O\left((a_{\max}\Upsilon\tilde{\lambda}_{\text{comm}}T)^{1+\frac{1}{p}} \log(1/\varepsilon)\, L_A + \lambda_B^2T^2\right).$$

□

6.4 Commutator Bound

Note that the decomposition for partially randomized circuits $\mathcal{S}$ has a different commutator factor $\tilde{\alpha}_{\text{comm}}$ (compared to $\mathcal{P}$ and α_{comm} from before). This introduces a new $\tilde{\lambda}_{\text{comm}}$ when analyzing the partially randomized product formulae extracted from Richardson, which is analogous to the λ_{comm} of the product formulae extracted from Richardson. We provide a bound for this new commutator factor $\tilde{\lambda}_{\text{comm}}$

Lemma 45. *Let $H = H_A + H_B$, where $H_A = \sum_{i\in A} H_i$ is the deterministic part and $H_B = \sum_{i\in B} H_i$ is the randomized part, so that $H = \sum_{i\in A\cup B} H_i$. Let $\tilde{\alpha}_{comm}$ be the nested commutator factor computed over a grouped decomposition of H_B, where each group is $\tilde{H}_{\tilde{\gamma}} = \sum_{i\in S_{\tilde{\gamma}}} H_i$ with $S_{\tilde{\gamma}} \subseteq B$ and $\bigcup_{\tilde{\gamma}} S_{\tilde{\gamma}} = B$ (as in equation (32)). Let $\tilde{\lambda}_{comm}$ be the resulting error bound from applying Lemma 11 using $\tilde{\alpha}_{comm}$. Let λ_{comm} be the error bound from applying Lemma 11 using the* ungrouped *decomposition $H = \sum_{i=1}^{\Gamma} H_i$ (with $\Gamma = |A| + |B|$). Then*

$$\tilde{\lambda}_{comm} \leq \lambda_{comm}.$$

Proof. We compare the nested commutator factors $\tilde{\alpha}^{(j)}_{\text{comm}}$ and $\alpha^{(j)}_{\text{comm}}$ for fixed order j.

Let $H' := \{H_i\}_{i\in A\cup B}$ be the ungrouped decomposition of the full Hamiltonian, and let the j-th order ungrouped commutator factor be

$$\alpha^{(j)}_{\text{comm}} := \sum_{i_1,\ldots,i_j\in A\cup B} \left\|[H_{i_1},[H_{i_2},\ldots,[H_{i_{j-1}},H_{i_j}]\ldots]]\right\|.$$

Now consider the grouped decomposition of H_B as $\tilde{H}_{\tilde{\gamma}} := \sum_{i\in S_{\tilde{\gamma}}} H_i$ for disjoint or overlapping subsets $S_{\tilde{\gamma}} \subseteq B$ such that $\bigcup_{\tilde{\gamma}} S_{\tilde{\gamma}} = B$.

Define the grouped commutator factor as:

$$\tilde{\alpha}^{(j)}_{\text{comm}} := \sum_{\tilde{\gamma}_1,\ldots,\tilde{\gamma}_j} \left\|[\tilde{H}_{\tilde{\gamma}_1},[\tilde{H}_{\tilde{\gamma}_2},\ldots,[\tilde{H}_{\tilde{\gamma}_{j-1}},\tilde{H}_{\tilde{\gamma}_j}]\ldots]]\right\|.$$

Now expand each $\tilde{H}_{\tilde{\gamma}}$ in terms of the H_i:

$$\tilde{H}_{\tilde{\gamma}_k} = \sum_{i_k\in S_{\tilde{\gamma}_k}} H_{i_k}, \quad \text{for each } k = 1,\ldots,j.$$

Then, by multilinearity of the nested commutator and the triangle inequality:

$$\begin{aligned}\left\|[\tilde{H}_{\tilde{\gamma}_1},\ldots,\tilde{H}_{\tilde{\gamma}_j}]\right\| &= \left\|\left[\sum_{i_1\in S_{\tilde{\gamma}_1}} H_{i_1},\left[\sum_{i_2\in S_{\tilde{\gamma}_2}} H_{i_2},\ldots,\left[\sum_{i_j\in S_{\tilde{\gamma}_j}} H_{i_j}\right]\ldots\right]\right]\right\| \\ &\le \sum_{i_1\in S_{\tilde{\gamma}_1}}\cdots\sum_{i_j\in S_{\tilde{\gamma}_j}} \left\|[H_{i_1},[H_{i_2},\ldots,[H_{i_{j-1}},H_{i_j}]\ldots]]\right\|.\end{aligned}$$

Summing over all group index tuples $(\tilde{\gamma}_1,\ldots,\tilde{\gamma}_j)$, we get:

$$\tilde{\alpha}^{(j)}_{\text{comm}} \le \sum_{i_1,\ldots,i_j\in B} \left\|[H_{i_1},[H_{i_2},\ldots,[H_{i_{j-1}},H_{i_j}]\ldots]]\right\| \le \alpha^{(j)}_{\text{comm}}.$$

That is, the grouped commutator factor is bounded above by the ungrouped one.

Since the error bound λ_{comm} is a monotonic function of the commutator factors (e.g., as in Lemma 11, the bound is linear or polynomial in $\alpha^{(j)}_{\text{comm}}$), it follows that $\tilde{\lambda}_{\text{comm}} \le \lambda_{\text{comm}}$ as required. □

7 Application: Phase estimation

In this section, we explore a method for phase estimation that involves computing the cumulative distribution function (CDF) of the eigenvalue spectrum of a Hamiltonian H. This CDF can be formulated using the Heaviside step function, which we approximate via a Fourier series expansion over the interval $[-\pi,\pi]$. Consequently, it is necessary to rescale the Hamiltonian such that its

spectrum lies within this interval. We define the normalized Hamiltonian as $\hat{H} := \kappa H$, where the scaling factor κ will be specified later. Throughout this section, we assume the existence of a known upper bound K on the spectral norm of H, meaning $K \geq \|H\|$.

The primary quantity of interest is the cumulative distribution function associated with H, given by

$$C(x) = (\hat{p} * \Theta)(x) = \mathrm{Tr}[\rho\, \Theta(xI - \kappa H)] = \sum_{x \geq \kappa E_i} \mathrm{Tr}[\rho \Pi_i]\,,$$

where $\Theta(x)$ is the Heaviside step function, $\hat{p}(x) = \sum_i \mathrm{Tr}[\rho\Pi_i]\, \delta(x - \kappa E_i)$ represents the probability density function corresponding to ρ and the normalized Hamiltonian $\hat{H}$, and $*$ denotes convolution. Here, Π_i is the projector onto the eigenspace associated with the i-th eigenvalue of H, arranged in increasing order. We also define $\eta = \mathrm{Tr}[\rho\Pi_0]$ as the overlap between the ground state and the ansatz state.

7.1 Heaviside Approximation and Approximate CDF Construction

To get these necessary coefficients and time signals, we first present a useful Fourier series approximation to the Heaviside function valid in the domain $[-\pi, \pi]$.

Lemma 46 (Heaviside Fourier series [WBC22b] Lemma 1)**.** *There exists a Fourier series* $\widetilde{\Theta}(x) = \sum_{j\in S} \widetilde{\Theta}_j e^{ijx}$ *with* $S := \{0\} \cup \{\pm(2j+1)\}_{j=0}^{d}$ *and maximum "time parameter" scaling as* $d = O(u^{-1} \log(\varepsilon_F^{-1}))$ *where* ε_F *is our Fourier error, which satisfies*

1. $|\widetilde{\Theta}(x) - \Theta(x)| \leq \varepsilon_F \quad \forall x \in [-\pi+u, -u] \cup [u, \pi-u]$, *where* $u \in [0, \pi/2]$ *serves as a resolution parameter.*
2. $-\varepsilon_F \leq |\widetilde{\Theta}(x)| \leq 1 + \varepsilon_F \quad \forall x \in \mathbb{R}$,
3. $\|\widetilde{\Theta}\|_1 := \sum_{j\in S_1} |\widetilde{\Theta}_j| = O(\log d)$.

We use this Fourier series to construct an approximate CDF which we denote as

$$\widetilde{C}(x) = (\hat{p} * \Theta)(x) = \sum_{j\in S} \widetilde{\Theta}_j \int_{-\pi}^{\pi} \hat{p}(y) e^{ij(x-y)} \mathrm{d}y = \sum_{j\in S} \widetilde{\Theta}_j e^{ijx}\, \mathrm{Tr}\left[\rho e^{-ij\kappa H}\right] = \mathrm{Tr}[\rho \widetilde{\Theta}(xI - \kappa H)]\,.$$

The following lemma ensures that this Fourier approximation gives a reasonable approximation to our desired cumulative distribution function.

Lemma 47 (Approximation to CDF from Fourier series, adapted from [WBC22b] Proposition 12)**.** *Take* $\widetilde{\Theta}(x)$ *in Lemma 46, and set* $\kappa = \frac{\pi-u}{2K}$. *Then, the quantity* $\widetilde{C}(x) = \mathrm{Tr}[\rho\widetilde{\Theta}(xI - \kappa H)]$ *satisfies*

$$C(x-u) - \varepsilon_F \leq \widetilde{C}(x) \leq C(x+u) + \varepsilon_F$$

We observe that the probability density function is defined over the interval $[-\kappa K, \kappa K]$, which lies within the range $[-\frac{1}{2}(\pi-u), \frac{1}{2}(\pi-u)]$. This choice of support is necessary to guarantee the error bounds established in equation (47). From this point onward, we fix the normalization factor as $\kappa = \frac{\pi-u}{2K}$ to ensure that this lemma remains applicable throughout the analysis.

7.2 Ground State Identification via Approximate CDF

Next, we demonstrate that the ground state energy can be characterized by a condition on the approximate cumulative distribution function (CDF), allowing us to use the approximate CDF as a tool to determine the ground state energy.

Lemma 48. *Suppose we have an algorithm to approximate $\widetilde{C}(x)$ to precision $\eta/8$ for some $0 < u < \pi/2$, $\varepsilon_F = \eta/8$ and for any $x \in [-\pi/2, \pi/2]$. Then we may determine whether*

$$C(x+u) > \eta/2 \quad \textit{or} \quad C(x-u) < \eta\,.$$

Suppose further that we locate an x^* that simultaneously satisfies both conditions. Then, we have that

$$|x^*/\kappa - E_0| \leq u/\kappa\,,$$

i.e., we have an additive estimate to the ground state energy E_0.

Proof. We use the algorithm to determine whether $\widetilde{C}(x) > 3\eta/4$ or not. Within the specified precision, we can guarantee either

$$\widetilde{C}(x) > (5/8)\eta \quad \textit{or} \quad \widetilde{C}(x) < (7/8)\eta\,,$$

which respectively implies either $C(x+u) > \eta/2$ or $C(x-u) < \eta$ respectively by use of Lemma 47, and we have satisfied the first claim. Now we suppose both conditions are simultaneously satisfied for some x^*. We recall that $\eta \leq \mathrm{Tr}[\Pi_0\rho]$ lower bounds the ground space overlap. We additionally note that the exact CDF $C(x)$ cannot take values in $(0, \mathrm{Tr}[\Pi_0\rho]) \supseteq (0, \eta)$ and so we have that

$$\begin{aligned} &C(x^*+u) \geq \eta, \quad C(x^*-u) = 0\,, \\ \implies &x^*+u \geq \kappa E_0\,, \quad x^*-u < \kappa E_0\,. \end{aligned}$$

Thus, $|x^* - \kappa E_0| \leq u$. □

Next we present an algorithm and lemma showing that the search algorithm can locate x^* using only a logarithmic number of queries to the approximate CDF $\widetilde{C}(x)$.

7.3 Binary Search for the Ground State Energy

Lemma 49 (Binary search with approximate CDF)**.** *Suppose that we have an algorithm $A(x, \widetilde{u}, \widetilde{\varepsilon}, \widetilde{\delta})$ which evaluates $\widetilde{C}(x)$ for any x to additive error $\widetilde{\varepsilon}$ and success probability at least $(1-\widetilde{\delta})$, for $\varepsilon_F = \eta/8$ and some resolution parameter $\widetilde{u} \in [0, \pi/2]$. Then, the ground state energy can be found to additive precision u/κ and success probability at least $(1-\delta)$ by running $A(x, 0.9u, \eta/8, \delta/L)$ at $L = O(\log(1/u))$ different values of x.*

Proof. The contents of [LT22, Section 5] prove this statement, but we present it here for completeness. We see from Lemma 48 that setting $\widetilde{\varepsilon} = \eta/8$ allows us to determine the criterion (48) (with some success probability to be later addressed and some resolution parameter). We first show how to use

this to search for x^*, before discussing how to tune the success probability. The search procedure operates by finding successively tighter upper and lower bounds to x^*. Set

$$x_{0,0} = -\pi/2, \quad x_{1,0} = \pi/2\,,$$

where we have that $C(x_{0,0}) < \eta$ and $C(x_{1,0}) > \eta/2$. We now specify an update rule to generate $(C(x_{0,\ell}), C(x_{1,\ell}))$ that have decreasing separation for increasing ℓ, but still satisfy $C(x_{0,\ell}) < \eta$ and $C(x_{1,\ell}) > \eta/2$ (recall that as $C(x)$ cannot take values in $(0, \eta)$, we have $x_{0,\ell} \leq x^* \leq x_{1,\ell}$). For all $\ell \geq 0$ until termination, first construct $x_\ell = (x_{0,\ell} + x_{1,\ell})/2$. Second, run $A(x_\ell, 0.9u, \eta/8, \widetilde{\delta})$ (note the prefactor 0.9 for the resolution parameter can be arbitrarily chosen to any number smaller than 1). Due to (48) this determines whether (i): $C(x_\ell + 0.9u) > \eta/2$ or (ii): $C(x_\ell - 0.9u) < \eta$. We now apply the following update rule based on the outcome

$$\begin{aligned} \text{(i)}: &\qquad x_{0,\ell+1} = x_{0,\ell}\,, &&\qquad x_{1,\ell+1} = x_\ell + 0.9u\,, \\ \text{(ii)}: &\qquad x_{0,\ell+1} = x_\ell - 0.9u\,, &&\qquad x_{1,\ell+1} = x_{1,\ell}\,, \end{aligned}$$

and it is simple to check that the conditions $C(x_{0,\ell}) < \eta$ and $C(x_{1,\ell}) > \eta/2$ are satisfied, and that the separation $x_{1,\ell} - x_{0,\ell}$ is decreasing with increasing ℓ so long as $x_{1,\ell} - x_{0,\ell} \geq 1.8u$. We terminate the procedure at step L when $x_{1,\ell} - x_{0,\ell} \leq 2u$, as this implies that $|x_L - \kappa E_0| \leq u$. Then, x_L/κ satisfies our desired approximation of the ground state energy E_0. One can check that the separation satisfies

$$x_{1,\ell} - x_{0,\ell} = \frac{\pi - 1.8u}{2^\ell} + 1.8u\,,$$

and thus $L = O(\log(1/u))$ evaluations of $A(x_\ell, 0.9u, \eta/8, \widetilde{\delta})$ are sufficient.

Finally, let us determine a sufficient success probability $\widetilde{\delta}$ for the algorithm that evaluates $\widetilde{C}(x)$. If each use of $A(x_\ell, 0.9u, \eta/8, \widetilde{\delta})$ fails with probability at most $\widetilde{\delta}$, then in L uses the failure probability is at most $L\widetilde{\delta}$. Our stated claim follows by requiring that this equals our desired overall failure probability $L\widetilde{\delta} = \delta$. □

Our last piece to construct a full phase estimation algorithm is to specify the algorithm $A(x, \widetilde{u}, \widetilde{\varepsilon}, \widetilde{\delta})$ to evaluate $\widetilde{C}(x)$. Observe that we evaluate $\widetilde{C}(x)$ statistically via our randomized quantum algorithm 1, where compilation is applied on the Heaviside function Θ. Through this algorithm we use the extrapolation strategy of Lemma 13, which uses extrapolation coefficients $\{b_k\}_k$ with accompanying product formulae $\{\mathrm{Tr}[\rho \mathcal{P}^{1/s_k}(s_k T)]\}_k$ using $\{1/s_k\}_k$ Trotter steps.

Algorithm 3 Randomized Estimation of Approximate CDF via Fourier–Heaviside Expansion

1: **Preprocessing:** Compute sampling distribution $\left\{\frac{|\widetilde{\Theta}_j b_k|}{\mathcal{Z}}\right\}$ for $j \in S$, $k = 1, \ldots, m$, where $\mathcal{Z} = \sum_{j,k} |\widetilde{\Theta}_j b_k|$
2: **for** $i = 1$ to C_{sample} **do**
3: Sample $(j', k') \sim$ distribution, $T \leftarrow \kappa \cdot j'$
4: Construct quantum circuit for $\mathcal{P}^{1/s_{k'}}(s_{k'}T)$
5: Prepare Hadamard tests to estimate real and imaginary parts:
6: $X_{\text{Re}}^{(k')} = \text{Re}[\text{Tr}(\rho \mathcal{P}^{1/s_{k'}}(s_{k'}T))]$, $X_{\text{Im}}^{(k')} = \text{Im}[\text{Tr}(\rho \mathcal{P}^{1/s_{k'}}(s_{k'}T))]$
7: Measure each circuit once; set $Z^{(i)} = X_{\text{Re}}^{(k')} + iX_{\text{Im}}^{(k')}$
8: $C^{(i)} \leftarrow \|\vec{b}\|_1 \cdot \|\widetilde{\Theta}\|_1 \cdot \text{sign}(b_{k'}\widetilde{\Theta}_{j'}) \cdot e^{ij'x} \cdot Z^{(i)}$
9: **end for**
10: **Return:** $\widetilde{C}(x) = \frac{1}{C_{\text{sample}}} \sum_{i=1}^{C_{\text{sample}}} C^{(i)}$

Similar to our randomized Algorithm 1, this algorithm is structured into two main parts: a quantum component (Step 6) followed by classical post-processing (Step 8). While we may want to evaluate the algorithm's output $A(x, \widetilde{u}, \widetilde{\varepsilon}, \widetilde{\delta})$ at multiple values of x, it is important to note that the quantum step does not need to be repeated for each x. Instead, we can run the quantum step once, record its outcomes, and then apply different classical post-processing (Step 8) corresponding to each desired x. As a result, obtaining outputs for L different values of x requires only a single round of quantum measurements, with all additional computation handled classically.

Lemma 50. *Algorithm 3 implements the procedure $A(x, \widetilde{u}, \widetilde{\varepsilon}, \widetilde{\delta})$, preparing a $\widetilde{\varepsilon}$-additive approximation to the approximate CDF $\widetilde{C}(x)$, with Fourier parameters $\widetilde{u}$ and $\varepsilon_F = \eta/8$, and overall success probability at least $1 - \widetilde{\delta}$.*

The algorithm uses the following resources:

$$C_{gate} = O\left(\Gamma\left(a_{\max}\Upsilon\lambda_{\text{comm}}\kappa\widetilde{u}^{-1}\log(\eta^{-1})\right)^{1+1/p} \cdot \log\left(\widetilde{\varepsilon}^{-1}\log\left(\widetilde{u}^{-1}\log(\eta^{-1})\right)\right)\right),$$

$$C_{sample} = O\left(\frac{\left(\log\log(\widetilde{\varepsilon}^{-1}) \cdot \log(\widetilde{u}^{-1}\log(\eta^{-1}))\right)^2}{\widetilde{\varepsilon}^2} \cdot \log\left(\frac{1}{\widetilde{\delta}}\right)\right),$$

where C_{sample} denotes the number of quantum circuit repetitions, and C_{gate} is the maximum depth of any individual circuit.

Proof. Algorithm 3 applies the randomized estimation primitive from Algorithm 1 to the smoothed Heaviside function $\widetilde{\Theta}$, which amounts to estimating

$$\widetilde{C}(x) = \text{Tr}\left[\rho\,\widetilde{\Theta}(xI - \kappa H)\right] = \text{Tr}[Zf(A)] \quad \text{with } Z = \rho,\ f(A) = \widetilde{\Theta}(xI - \kappa H).$$

We apply Lemma 23 using the following parameters:

- $\|Z\|_1 = 1$ (since ρ is a density matrix),

- Fourier parameters $\widetilde{u}$ and $\varepsilon_F = \eta/8$,
- target additive error $\widetilde{\varepsilon}$.

From Lemma 46, the Fourier approximation of $\widetilde{\Theta}$ has:

$$d = O(\widetilde{u}^{-1}\log(1/\eta)), \quad c = \|\widetilde{\Theta}\|_1 = O(\log(\widetilde{u}^{-1}\log(1/\eta))), \quad T = \kappa d = O(\kappa\widetilde{u}^{-1}\log(1/\eta)).$$

Substituting into Lemma 23 yields:

$$C_{\text{sample}} = O\left(\frac{\left(\log\log(\widetilde{\varepsilon}^{-1})\cdot\log(\widetilde{u}^{-1}\log(\eta^{-1}))\right)^2}{\widetilde{\varepsilon}^2}\cdot\log\left(\frac{1}{\widetilde{\delta}}\right)\right),$$

$$C_{\text{gate}} = O\left(\Gamma\left(a_{\max}\Upsilon\lambda_{\text{comm}}\kappa\widetilde{u}^{-1}\log(\eta^{-1})\right)^{1+1/p}\cdot\log\left(\widetilde{\varepsilon}^{-1}\log\left(\widetilde{u}^{-1}\log(\eta^{-1})\right)\right)\right).$$

□

The lemma establishes the correctness and resource costs of the approximate CDF estimation procedure implemented by Algorithm 3. We now leverage this subroutine within a binary search framework to efficiently estimate the ground state energy, as detailed in Algorithm 4.

Algorithm 4 Binary Search for Ground State Energy Using Approximate CDF

Input: precision u/κ, resolution $u \in [0, \pi/2]$, threshold η, failure probability δ, subroutine $A(x, \widetilde{u}, \widetilde{\varepsilon}, \widetilde{\delta})$ (Alg. 3)
Output: estimate $\widehat{E}_0$ such that $|\widehat{E}_0 - E_0| \le u/\kappa$ with probability $\ge 1 - \delta$
Set $\widetilde{\varepsilon} \leftarrow \eta/8$, $\widetilde{u} \leftarrow 0.9u$, $L \leftarrow \lceil \log_2(\pi/u) \rceil$, $\widetilde{\delta} \leftarrow \delta/L$
Initialize: $x_{0,0} \leftarrow -\pi/2$; $\quad x_{1,0} \leftarrow \pi/2$ ▷ $C(x_{0,0}) < \eta$, $C(x_{1,0}) > \eta/2$
for $\ell = 0$ to $L - 1$ **do**
    Set $x_\ell \leftarrow \frac{x_{0,\ell} + x_{1,\ell}}{2}$
    Run $A(x_\ell, \widetilde{u}, \widetilde{\varepsilon}, \widetilde{\delta})$ to check:
    - (i) $C(x_\ell + \widetilde{u}) > \eta/2$ or (ii) $C(x_\ell - \widetilde{u}) < \eta$

    Update:

$$\begin{cases} x_{0,\ell+1} \leftarrow x_{0,\ell}; \quad x_{1,\ell+1} \leftarrow x_\ell + \widetilde{u} & \text{if (i)} \\ x_{0,\ell+1} \leftarrow x_\ell - \widetilde{u}; \quad x_{1,\ell+1} \leftarrow x_{1,\ell} & \text{if (ii)} \end{cases}$$

    if $x_{1,\ell+1} - x_{0,\ell+1} \le 2u$ **then break**
    end if
end for
Set $x_L \leftarrow \frac{x_{0,L} + x_{1,L}}{2}$;
return $\widehat{E}_0 \leftarrow x_L/\kappa$

Theorem 51. *The ground state energy estimation problem can be solved to precision ε and with*

success probability at least $(1-\delta)$ *using the above algorithm with resources*

$$C_{sample} = O\left(\frac{\left(\log\log(\eta^{-1})\cdot\log(K\varepsilon^{-1}\log(\eta^{-1}))\right)^2}{\eta^2}\log\left(\frac{\log(K\varepsilon^{-1})}{\delta}\right)\right)$$

$$C_{gate} = O\left(\Gamma\left(a_{\max}\Upsilon\lambda_{\text{comm}}\varepsilon^{-1}\log(\eta^{-1})\right)\right)^{(1+1/p)}\log(\eta^{-1}\log(K\varepsilon^{-1}\log(\eta^{-1})))\Bigg)$$

Proof. We use the search algorithm outlined in Lemma 49 and descibed in pseudocode with Algorithm 4 which makes calls to an algorithm $A(x,\widetilde{u},\widetilde{\varepsilon},\widetilde{\delta})$ which approximately returns the approximate CDF, and use Lemma 50 to instantiate $A(x,\widetilde{u},\widetilde{\varepsilon},\widetilde{\delta})$. Set $\widetilde{\varepsilon}=\eta/8$, $\widetilde{u}=0.9u$, $\widetilde{\delta}=\delta/L$ with $L=O(\log(1/u))$, as discussed in Lemma 49. Further, recall from Lemma 49 that to reach additive precision ε in the ground state energy we set $u=\kappa\varepsilon$, and that our CDF is defined such that $\kappa=O(K^{-1})$, where K is an upper bound on the spectral norm of H (see Lemma 47). □

8 Application: Green's functions

In this section, we apply our algorithm to the evaluation of Green's functions in the context of many-body physics. Green's functions capture key dynamical and spectral properties of quantum systems, and their accurate estimation provides insights into physical observables such as particle propagation, kinetic energy, and spectral densities.

8.1 Obtaining the Resolvent from the Fourier Representation

As in the previous application, to evaluate Green's functions, we first consider their Fourier representation. This gives us the desired resolvent operator, which we then need to estimate.

Lemma 52 (Resolvent Representation via Fourier–Laplace Transform ([KDW21]))**.** *Let* $G^R_{jj_0}(t)$ *denote the time-domain retarded Green's functiFon of a quantum system with N particles and ground state energy* E_0^N*. Then, its frequency-domain representation is obtained via the Fourier–Laplace integral transform with exponential convergence factor* $e^{-\Gamma_{broad}t}$ $(\Gamma_{broad}>0)$ *as:*

$$G^R_{jj_0}(\omega) = \int_0^\infty dt\, G^R_{jj_0}(t)e^{i(\omega+i\Gamma_{broad})t},$$

which evaluates to:

$$\begin{aligned}G^R_{jj_0}(\omega) = {}& \langle\psi_0^N|\,\hat{c}_{j_0}R(\omega^+ + i\Gamma_{broad},\hat{H}_0)\hat{c}_j^\dagger\,|\psi_0^N\rangle \\ & + \langle\psi_0^N|\,\hat{c}_j^\dagger R(\omega^- + i\Gamma_{broad},-\hat{H}_0)\hat{c}_{j_0}\,|\psi_0^N\rangle\,,\end{aligned}$$

where:

- $\omega^\pm = \omega \pm \mu(\hat{N}-N)$,
- $\hat{H}_0 = \hat{H} - E_0^N$,

- $R(\omega + i\Gamma_{broad}, \hat{H})$ *is the resolvent operator, defined as:*

$$R(\omega + i\Gamma_{broad}, \hat{H}) = -i \int_0^\infty dt\, e^{i(\omega + i\Gamma_{broad} - \hat{H})t}. \tag{33}$$

Evaluating Eq. (33) analytically yields the standard form of the resolvent operator:

$$R(\omega + i\Gamma_{broad}, \hat{H}) = (\omega + i\Gamma_{broad} - \hat{H})^{-1}.$$

This shows that the frequency-domain Green's function can be expressed in terms of a resolvent operator, which is a shifted inverse of the Hamiltonian.

8.2 Approximation of the Resolvent

Back to estimating Green's functions, we need to approximate the resolvent operator. Using the Linear Combination of Unitaries (LCU) technique [CW12], the resolvent

$$R(\omega + i\Gamma_{\text{broad}}, \hat{H}) = (\omega + i\Gamma_{\text{broad}} - \hat{H})^{-1}$$

can be approximated with additive error ε using

$$h(\omega + i\Gamma_{\text{broad}}, \hat{H}) = -i \sum_{k=0}^{N_c} \Delta t\, e^{i(\omega + i\Gamma_{\text{broad}} - \hat{H})k\Delta t}.$$

Lemma 53 (LCU Approximation of the Resolvent)**.** *[CW12] Let $\hat{H}$ have spectrum $\sigma(\hat{H}) \subseteq [0, 1]$. For any $\omega \in \sigma(\hat{H})$ and artificial broadening $\Gamma_{broad} > 0$, the above LCU sum satisfies*

$$\|R(\omega + i\Gamma_{broad}, \hat{H}) - h(\omega + i\Gamma_{broad}, \hat{H})\| \leq \varepsilon$$

provided

$$N_c = O\Big(\frac{1}{\Gamma_{broad}\varepsilon} \log \frac{1}{\Gamma_{broad}\varepsilon}\Big), \qquad \Delta t = \min\Big\{\frac{\varepsilon}{2}, \frac{3}{\|\hat{H}\|}\Big\} = O(\varepsilon).$$

Where $N_c = t_c/\Delta t$ is the number of discrete steps. Implementing this approximation requires $O\big(\frac{1}{\Gamma_{broad}^2} \log \frac{1}{\Gamma_{broad}\varepsilon}\big)$ queries to a time-evolution oracle for $\hat{H}$ and $\log N_c$ ancilla qubits.

To compute the Fourier coefficients and corresponding time signals c and T required by Algorithm 1, we use the following lemma, which provides bounds on truncation and discretization of the Fourier-Laplace integral defining the resolvent.

Lemma 54 (Truncation Bound for Discretized Resolvent)**.** *Let $\Gamma_{broad} > 0$ and $\varepsilon > 0$. Truncating the Fourier-Laplace integral*

$$R(\omega + i\Gamma_{broad}, \hat{H}) = \int_0^\infty dt\, e^{i(\omega - \hat{H})t} e^{-\Gamma_{broad}t}$$

at $t_c = N_c\Delta t$ and discretizing with step size Δt gives

$$h(\omega + i\Gamma_{broad}, \hat{H}) = \sum_{k=0}^{N_c} \Delta t\, e^{-\Gamma_{broad} k\Delta t} e^{-i(\hat{H}-\omega)k\Delta t}.$$

To ensure the truncation error

$$\varepsilon_t = \|R(\omega + i\Gamma_{broad}, \hat{H}) - h_\infty(\omega + i\Gamma_{broad}, \hat{H})\| < \frac{\varepsilon}{2},$$

it suffices to choose

$$t_c = N_c\Delta t = \frac{1}{\Gamma_{broad}} \log \frac{2}{\Gamma_{broad}\varepsilon}, \qquad N_c = \frac{t_c}{\Delta t} = \frac{1}{\Gamma_{broad}\Delta t} \log \frac{2}{\Gamma_{broad}\varepsilon}.$$

Proof. The truncation error comes from neglecting the tail beyond t_c:

$$\varepsilon_t = \left\| \int_{t_c}^{\infty} dt\, e^{i(\omega-\hat{H})t} e^{-\Gamma_{\text{broad}} t} \right\| \leq \int_{t_c}^{\infty} dt\, e^{-\Gamma_{\text{broad}} t} = \frac{e^{-\Gamma_{\text{broad}} t_c}}{\Gamma_{\text{broad}}}.$$

Requiring $\varepsilon_t < \varepsilon/2$ gives

$$\frac{e^{-\Gamma t_c}}{\Gamma_{\text{broad}}} < \frac{\varepsilon}{2} \quad \Longrightarrow \quad e^{-\Gamma_{\text{broad}} t_c} < \frac{\Gamma_{\text{broad}}\varepsilon}{2} \quad \Longrightarrow \quad t_c = \frac{1}{\Gamma_{\text{broad}}} \log \frac{2}{\Gamma_{\text{broad}}\varepsilon}.$$

The number of discrete steps is then $N_c = t_c/\Delta t$, yielding

$$N_c = \frac{1}{\Gamma_{\text{broad}}\Delta t} \log \frac{2}{\Gamma_{\text{broad}}\varepsilon}.$$

This guarantees the truncation error is within $\varepsilon/2$. □

Finally, we apply our randomized primitive to estimate the resolvent operator for the Green's function.

Theorem 55 (Estimating the Resolvent via Randomized LCU)**.** *Assuming that we can prepare a ground state, the resolvent operator*

$$R(\omega + i\Gamma_{broad}, \hat{H}) = (\omega + i\Gamma_{broad} - \hat{H})^{-1}$$

can be estimated to additive error ε using Algorithm 1 with the following complexities:

$$C_{gate} = O\left(\log \frac{1}{\Gamma_{broad}\varepsilon} \cdot \left(a_{max}\, \Upsilon\, \lambda_{comm} \cdot \frac{1}{\Gamma_{broad}} \log \frac{1}{\Gamma_{broad}\varepsilon} \right)^{1+\frac{1}{p}} \right),$$

$$C_{sample} = O\left(\frac{\left(\frac{1}{\Gamma_{broad}} \log \frac{1}{\Gamma_{broad}\varepsilon} \right)^2 (\log\log \frac{1}{\varepsilon})^2}{\varepsilon^2} \cdot \log \frac{1}{\delta} \right).$$

Proof. From Lemma 52, the resolvent can be approximated as

$$h(\omega + i\Gamma_{\text{broad}}, \hat{H}) = -i \sum_{k=0}^{N_c} \Delta t \, e^{i(\omega + i\Gamma_{\text{broad}} - \hat{H})k\Delta t}$$

with additive error ε. The corresponding Fourier coefficients and time signals are bounded by

$$c = N_c \Delta t, \quad t_{\max} = t_c,$$

where Lemma 54 gives

$$t_c = \frac{1}{\Gamma_{\text{broad}}} \log \frac{2}{\Gamma_{\text{broad}} \varepsilon}, \quad N_c = \frac{1}{\Gamma_{\text{broad}} \Delta t} \log \frac{2}{\Gamma_{\text{broad}} \varepsilon}.$$

Thus, we have

$$c = O\left(\frac{1}{\Gamma_{\text{broad}}} \log \frac{1}{\Gamma_{\text{broad}} \varepsilon}\right), \quad t_{\max} = O\left(\frac{1}{\Gamma_{\text{broad}}} \log \frac{1}{\Gamma_{\text{broad}} \varepsilon}\right).$$

Applying these parameters in Algorithm 1 yields the stated gate and sample complexities:

$$C_{\text{gate}} = O\left(\log \frac{1}{\Gamma_{\text{broad}} \varepsilon} \cdot \left(a_{\max} \Upsilon \lambda_{\text{comm}} \cdot \frac{1}{\Gamma_{\text{broad}}} \log \frac{1}{\Gamma_{\text{broad}} \varepsilon} \right)^{1+\frac{1}{p}} \right),$$

$$C_{\text{sample}} = O\left(\frac{\left(\frac{1}{\Gamma_{\text{broad}}} \log \frac{1}{\Gamma_{\text{broad}} \varepsilon} \right)^2 (\log \log \frac{1}{\varepsilon})^2}{\varepsilon^2} \cdot \log \frac{1}{\delta} \right).$$

□

9 Application: Time-Evolved States

Theorem 29 directly gives an application for probing the distributions of time-evolved states, that is, by considering the function $f(H) = \exp(-iHt)$. Here the Fourier parameters are trivial and the result can be directly stated.

Theorem 56 (Distribution recovery for time-evolved states)**.** *Consider a Hamiltonian of the form in Definition 1 and denote its time-evolved distribution on accessible input state $|\psi\rangle$ as the vector $\vec{p}$ with entries $p_i := |\langle i| e^{-iHT} |\psi\rangle|^2$. We give an algorithm to return $\vec{v}$ such that $\|\vec{v} - \vec{p}\| \leq \varepsilon$ with success probability at least $(1 - \delta)$ and*

- ***Gate complexity (per sample):***

$$C_{\text{gate}} = O\left(\Gamma \left(a_{\max} \Upsilon \lambda_{\text{comm}} T \right)^{1+\frac{1}{p}} \log(1/\varepsilon) \right),$$

- ***Sample complexity:***

$$C_{\text{sample}} = O\left(\frac{(\log \log(1/\varepsilon))^4}{\varepsilon^2} \cdot \log\left(\frac{1}{\delta}\right) \right),$$

- ***Classical preprocessing time:***

$$C_{\text{pre}} = O(\log(1/\varepsilon)).$$

Further, the algorithm uses one ancillary qubit.

References

[BC12] Dominic W Berry and Andrew M Childs. 'Black-box hamiltonian simulation and unitary implementation'. In: *Quantum Information and Computation* 12.1-2 (2012), pp. 29–62.

[Cam19] Earl Campbell. 'Random Compiler for Fast Hamiltonian Simulation'. In: *Phys. Rev. Lett.* 123.7 (Aug. 2019). arXiv:1811.08017.

[Cam21] Earl T Campbell. 'Early fault-tolerant simulations of the Hubbard model'. In: *Quantum Science and Technology* 7.1 (2021). arXiv:2012.09238, p. 015007.

[Cha+25] Shantanav Chakraborty, Soumyabrata Hazra, Tongyang Li, Changpeng Shao, Xinzhao Wang and Yuxin Zhang. 'Quantum singular value transformation without block encodings: Near-optimal complexity with minimal ancilla'. arXiv:2504.02385. 2025.

[Chi+21a] Andrew M Childs, Yuan Su, Minh C Tran, Nathan Wiebe and Shuchen Zhu. 'Theory of trotter error with commutator scaling'. In: *Physical Review X* 11.1 (2021), p. 011020.

[Chi+21b] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe and Shuchen Zhu. 'Theory of Trotter Error with Commutator Scaling'. In: *Phys. Rev. X* 11.1 (Feb. 2021). arXiv:1912.08854.

[CW12] Andrew M Childs and Nathan Wiebe. 'Hamiltonian simulation using linear combinations of unitary operations'. In: *Quantum Information & Computation* 12.11-12 (2012), pp. 901–924.

[Dal+23] Alexander M Dalzell et al. 'Quantum algorithms: A survey of applications and end-to-end complexities'. arXiv:2310.03011. 2023.

[DLT22] Yulong Dong, Lin Lin and Yu Tong. 'Ground-state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices'. In: *PRX Quantum* 3.4 (2022). arXiv:2204.05955, p. 040305.

[Gün+25] Jakob Günther, Freek Witteveen, Alexander Schmidhuber, Marek Miller, Matthias Christandl and Aram Harrow. 'Phase estimation with partially randomized time evolution'. arXiv:2503.05647. 2025.

[HHL09] Aram W. Harrow, Avinatan Hassidim and Seth Lloyd. 'Quantum algorithm for linear systems of equations'. In: *Phys. Rev. Lett.* 103.15 (2009). arXiv:0811.3171, p. 150502.

[HHZ19] Markus Heyl, Philipp Hauke and Peter Zoller. 'Quantum localization bounds Trotter errors in digital quantum simulation'. In: *Science advances* 5.4 (2019). arXiv:1806.11123, eaau8342.

[KDW21] Trevor Keen, Eugene Dumitrescu and Yan Wang. *Quantum Algorithms for Ground-State Preparation and Green's Function Calculation*. 2021. arXiv: 2112.05731 [quant-ph].

[Kis+25] Oriel Kiss, Utkarsh Azad, Borja Requena, Alessandro Roggero, David Wakeham and Juan Miguel Arrazola. 'Early fault-tolerant quantum algorithms in practice: Application to ground-state energy estimation'. In: *Quantum* 9 (2025). arXiv:2405.03754, p. 1682.

[LC17] Guang Hao Low and Isaac L. Chuang. 'Optimal Hamiltonian Simulation by Quantum Signal Processing'. In: *Phys. Rev. Lett.* 118.1 (2017). arXiv:1606.02685, p. 010501.

[LC19] Guang Hao Low and Isaac L Chuang. 'Hamiltonian simulation by qubitization'. In: *Quantum* 3 (2019), p. 163.

[LKW19] Guang Hao Low, Vadym Kliuchnikov and Nathan Wiebe. 'Well-conditioned multi-product Hamiltonian simulation'. arXiv:1907.11679. 2019.

[Llo96] Seth Lloyd. 'Universal Quantum Simulators'. In: *Science* 273.5278 (1996), pp. 1073–1078.

[Low+23] Guang Hao Low, Yuan Su, Yu Tong and Minh C. Tran. 'Complexity of Implementing Trotter Steps'. In: *PRX Quantum* 4.2 (May 2023).

[LT22] Lin Lin and Yu Tong. 'Heisenberg-Limited Ground-State Energy Estimation for Early Fault-Tolerant Quantum Computers'. In: *PRX Quantum* 3.1 (2022). arXiv:2102.11340, p. 010318.

[MCS22] Sam McArdle, Earl Campbell and Yuan Su. 'Exploiting fermion number in factorized decompositions of the electronic structure Hamiltonian'. In: *Phys. Rev. A* 105 (1 Jan. 2022). arXiv:2107.07238, p. 012403.

[Miz26] Kaoru Mizuta. 'On the commutator scaling in Hamiltonian simulation with multi-product formulas'. In: *Quantum* 10 (2026). arXiv:2507.06557, p. 1974.

[SHC21] Yuan Su, Hsin Yuan Huang and Earl T. Campbell. 'Nearly tight Trotterization of interacting electrons'. In: *Quantum* 5.1 (2021). arXiv:2012.09194, pp. 1–58. arXiv: 2012.09194.

[Tra+20] Minh C Tran, Su-Kuan Chu, Yuan Su, Andrew M Childs and Alexey V Gorshkov. 'Destructive error interference in product-formula lattice simulation'. In: *Physical review letters* 124.22 (2020), p. 220502.

[Wan+23] Guoming Wang, Daniel Stilck França, Ruizhe Zhang, Shuchen Zhu and Peter D Johnson. 'Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision'. In: *Quantum* 7 (2023). arXiv:2209.06811, p. 1167.

[Wan+25] Guoming Wang, Daniel Stilck França, Gumaro Rendon and Peter D Johnson. 'Efficient ground-state-energy estimation and certification on early fault-tolerant quantum computers'. In: *Physical Review A* 111.1 (2025). arXiv:2304.09827, p. 012426.

[WBC22a] Kianna Wan, Mario Berta and Earl T Campbell. 'Randomized quantum algorithm for statistical phase estimation'. In: *Physical Review Letters* 129.3 (2022), p. 030503.

[WBC22b] Kianna Wan, Mario Berta and Earl T. Campbell. 'Randomized Quantum Algorithm for Statistical Phase Estimation'. In: *Phys. Rev. Lett.* 129.3 (2022). arXiv:2110.12071, p. 030503.

[WMB24] Samson Wang, Sam McArdle and Mario Berta. 'Qubit-efficient randomized quantum algorithms for linear algebra'. In: *PRX Quantum* 5.2 (2024). arXiv:2302.01873, p. 020324.

[WW24a] James D Watson and Jacob Watkins. 'Exponentially Reduced Circuit Depths Using Trotter Error Mitigation'. In: *arXiv preprint arXiv:2408.14385* (2024).

[WW24b] James D Watson and Jacob Watkins. 'Exponentially Reduced Circuit Depths Using Trotter Error Mitigation'. arXiv:2408.14385. 2024.