



This is a repository copy of *Tailoring term truncations for electronic structure calculations using a linear combination of unitaries*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/187084/>

Version: Published Version

---

**Article:**

Meister, R., Benjamin, S.C. and Campbell, E.T. (2022) Tailoring term truncations for electronic structure calculations using a linear combination of unitaries. *Quantum*, 6. 637. ISSN 2521-327X

<https://doi.org/10.22331/q-2022-02-02-637>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:  
<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Tailoring Term Truncations for Electronic Structure Calculations Using a Linear Combination of Unitaries

Richard Meister<sup>1</sup>, Simon C. Benjamin<sup>1</sup>, and Earl T. Campbell<sup>2,3</sup>

<sup>1</sup>Department of Materials, University of Oxford, Oxford OX1 3PH, United Kingdom

<sup>2</sup>Department of Physics and Astronomy, University of Sheffield, Sheffield S3 7RH, United Kingdom

<sup>3</sup>AWS Center for Quantum Computing, Pasadena, CA 91125, USA

**A highly anticipated use of quantum computers is the simulation of complex quantum systems including molecules and other many-body systems. One promising method involves directly applying a linear combination of unitaries (LCU) to approximate a Taylor series by truncating after some order. Here we present an adaptation of that method, optimized for Hamiltonians with terms of widely varying magnitude, as is commonly the case in electronic structure calculations. We show that it is more efficient to apply LCU using a truncation that retains larger magnitude terms as determined by an iterative procedure. We obtain bounds on the simulation error for this generalized truncated Taylor method, and for a range of molecular simulations, we report these bounds as well as exact numerical results. We find that our adaptive method can typically improve the simulation accuracy by an order of magnitude, for a given circuit depth.**

## 1 Introduction

One of the most promising applications of quantum computers is the efficient simulation of quantum systems [1], including those that arise in quantum chemistry. Following the first concepts for such simulations [2, 3], there have been numerous proposed algorithms to simulate these systems using quantum computers [4–12], often with variations of Trotter-Suzuki product formulas [13, 14]. These methods usually approximate the time evolution operator by sequentially evolving the terms in the Hamiltonian individually. Through extensive study, the required gate count was reduced substantially over time [15–20]. However, the scaling of the inverse simulation error of such product formulas is polynomial in the circuit gate count.

An alternative is available through the technique of linear combinations of unitaries (LCU). Here, in contrast to the product formula approaches, one derives a quantum circuit that directly applies a sum of unitaries, allowing for a much greater variety of accessible operators. A key enhancement was the replacement of a probabilistic step in the original scheme [21] with

a near-deterministic process based on oblivious amplitude amplification [22].

The LCU method gave rise to a number of implementations for Hamiltonian simulation. The approach in Ref. [23] uses linear combinations of product formulas, taking advantage of commuting terms in the Hamiltonian – like pure product formulas – while improving the complexity scaling with inverse error to be only poly-logarithmic using LCU. In Ref. [24] it is applied to enhance the scaling with the error in Szegedy quantum walks while retaining their advantage for sparse Hamiltonians. Extensions of this approach are quantum signal processing [25, 26] and qubitization [27], of which variants specifically for quantum chemistry exist [28].

One of the most direct uses of LCU is presented in [29], where the time evolution operator is approximated by truncating its Taylor expansion at some appropriate order. This results in exponentially better scaling of the complexity with inverse error than for product formulas.

The aforementioned methods of qubitization and quantum signal processing have been shown to exhibit even better scaling for many types of Hamiltonians [27, 28, 30, 31]. However, there are instances where they are less suited, one prominent example being for simulating time-dependent Hamiltonians. Even for intrinsically time-independent cases, introducing a time dependence by transforming to a rotating frame can be beneficial if the Hamiltonian is diagonally dominant. In contrast to qubitization and quantum signal processing, the approach of the truncated Taylor series in [29] can be applied to such time-dependent cases with reasonable overhead, as shown in Refs. [32, 33], making it very relevant for such instances.

In this work, we present a variant of the truncated Taylor scheme [29] that includes terms by weight rather than order. By doing so, we try to exploit the fact that the Hamiltonians of some quantum mechanical systems – especially those of electrons in molecules – have terms whose magnitudes vary considerably. This suggests that some (large) terms should be included to higher orders than other (small) terms. Our variant implements just that, while respecting the efficient circuit implementation of [29] and subsequent improvements to SELECT and PREPARE subroutines [31, 34].

Our algorithm starts from an empty expansion and iteratively adds terms that facilitate the largest decline of the error bound for one additional gate. This greedy method leads to a more rapid reduction of the error in the very early stage of the construction when applied to electronic structure Hamiltonians. At a later stage, the rate of convergence becomes roughly equal to the original method, maintaining an approximately constant factor advantage in the error for the investigated molecules. Therefore, the asymptotic behavior is equivalent for both methods, but we accomplish a constant improvement. We find that the error of our modified scheme is typically one order of magnitude lower than in the original method at the same gate cost. For a fixed error magnitude, this results in reducing the circuit depth by roughly one full order of the expansion.

The rest of the paper is structured as follows. Section 2 contains a detailed description of our modified method adapted from [29]. In Section 3, we present results for error bounds as well as numerically evaluated errors for a variety of electronic structure Hamiltonians. Lastly, Section 4 concludes the paper and gives an outlook to possible further work.

## 2 Truncated Taylor series

Our method is closely related to the approach presented by Berry et al. [29]. We will give a detailed description of our modified method, which at the same time serves as a summary of [29].

### 2.1 Linear combination of unitaries

The protocol is based on a method of adding unitaries with the help of ancilla qubits [21]. We start from a Hamiltonian of the form

$$H = \sum_{\ell=0}^{L-1} \alpha_{\ell} h_{\ell}, \quad (1)$$

where  $\alpha_{\ell}$  are real positive scalars<sup>1</sup> and  $h_{\ell}$  are unitaries for which implementations on a quantum computer exist. Without loss of generality, we assume the terms are sorted by magnitude, i.e.  $\alpha_{\ell+1} \geq \alpha_{\ell}$ . The approach also used in [29] is to implement an approximation to the corresponding time evolution operator

$$U(t) = e^{-iHt} \quad (2)$$

with a Taylor series. Taking  $t$  to be sufficiently small, the series representation of  $U(t)$  can be approximated by the sum

$$U_{\mathbf{L}}(t) := \mathbb{1} + \sum_{k=1}^{\infty} \frac{(-it)^k}{k!} \prod_{j=1}^k \left( \sum_{\ell_j=0}^{L_j-1} \alpha_{\ell_j} h_{\ell_j} \right) \quad (3)$$

<sup>1</sup>Phases can always be pushed into the operators  $h_{\ell}$ .

where  $\mathbf{L}$  is a vector of  $L_k$  with  $0 \leq L_k \leq L$  and  $k \in \mathbb{N}^+$ , meaning the individual sums in the product only contain the  $L_k$  largest terms of  $H$ . This is the main difference to [29], where the series is truncated at some appropriate order  $n$ , which yields

$$U_n(t) := \mathbb{1} + \sum_{k=1}^n \frac{(-it)^k}{k!} \prod_{j=1}^k \left( \sum_{\ell_j=0}^{L_j-1} \alpha_{\ell_j} h_{\ell_j} \right). \quad (4)$$

Equation (4) is a special case of Eq. (3), where all orders up to  $n$  are added in full.<sup>2</sup> Our modified version of the sum includes some orders only partially, giving greater control over the total gate count and allowing for quicker convergence of the error bounds.

The magnitude of the time step  $t$  will be a fixed value restricted by the method. Longer times  $\tau = rt$ , can be simulated by applying  $U_{\mathbf{L}}^r$ . However, most of this paper will focus on the implementation of a single time step.

To keep the notation simple, the products of the coefficients  $\alpha_{\ell}$  with  $t^k/k!$  are gathered into new variables  $\beta_j$ , and all products of the unitaries  $h_{\ell}$  together with  $(-i)^k$  are collected into operators  $V_j$ , with a newly introduced label  $j$  numbering all terms in the sum. Note that even if different products of  $h_{\ell}$  yield identical operators, they are treated as separate  $V_j$ , each with a corresponding weight  $\beta_j$ . By construction, all  $\beta_j$  are also real and positive. Thus, Eq. (3) becomes

$$U_{\mathbf{L}} = \sum_{j=0}^{m-1} \beta_j V_j \quad (5)$$

where the time-dependence of  $U_{\mathbf{L}}$  and  $\beta_j$  is not explicitly denoted, and the total number of terms  $m$  implicitly depends on  $\mathbf{L}$ .

In order to apply  $U_{\mathbf{L}}$  to a state  $|\psi\rangle$ , we define the unitary operators  $\mathcal{P}(t)$  and  $\mathcal{S}$  (PREPARE and SELECT) in accordance with [29]. The PREPARE operator  $\mathcal{P}$ , whose time dependence we will make implicit from here on, maps the  $|0\rangle$  state of a set of ancilla qubits to the weighted superposition

$$\mathcal{P}|0\rangle := \frac{1}{\sqrt{s_{\mathbf{L}}}} \sum_{j=0}^{m-1} \sqrt{\beta_j} |j\rangle \quad (6)$$

with the implicitly  $t$ -dependent normalization constant

$$s_{\mathbf{L}} := \sum_{j=0}^{m-1} \beta_j. \quad (7)$$

The SELECT operator  $\mathcal{S}$  acts on a state  $|\psi\rangle$  with the operator  $V_j$ , where  $j$  is the state of the ancilla introduced above. So its action on a tensor state of  $|\psi\rangle$  with the ancilla  $|j\rangle$  is

$$\mathcal{S}|j\rangle|\psi\rangle := |j\rangle V_j |\psi\rangle. \quad (8)$$

<sup>2</sup>For all quantities with an  $\mathbf{L}$  subscript we will alternatively replace it with  $n$  to mean an  $\mathbf{L}$  where  $L_k = L$  for  $k \leq n$  and  $L_k = 0$  for  $k > n$ .

Given these two operators  $\mathcal{P}$  and  $\mathcal{S}$ , we proceed analogously to [29] by introducing a new operator

$$\mathcal{W} := (\mathcal{P}^\dagger \otimes \mathbb{1}) \mathcal{S} (\mathcal{P} \otimes \mathbb{1}) \quad (9)$$

which has the effect

$$\mathcal{W}|0\rangle|\psi\rangle = \frac{1}{s_{\mathbf{L}}}|0\rangle U_{\mathbf{L}}|\psi\rangle + N|0^\perp, \Phi\rangle \quad (10)$$

where  $N$  is the appropriate constant for the state to be normalized, and  $|0^\perp, \Phi\rangle$  is a garbage state whose ancilla part has no overlap with the ancillary  $|0\rangle$  state.

## 2.2 Oblivious amplitude amplification

The naïve method for obtaining  $U_{\mathbf{L}}|\psi\rangle$  would be to measure the ancilla of  $\mathcal{W}|0\rangle|\psi\rangle$ , see Eq. (10), and post-select for the ancilla  $|0\rangle$  state. However, since  $s_{\mathbf{L}}$  increases with  $t$ , the success probability for large  $t$  diminishes. Additionally,  $t$  is always subject to convergence of Eq. (3). Due to the postselection, dividing the total  $t$  into smaller segments and repeating the process multiple times would also suppress the total success probability.

One way around this problem also used in [29] is the so-called oblivious amplitude amplification. As detailed in Lemma 1 in Appendix B, and references therein, if  $U_{\mathbf{L}}$  were unitary and  $s_{\mathbf{L}} = 2$ , the amplification operator

$$\mathcal{Q} := -\mathcal{W}R\mathcal{W}^\dagger R \quad (11)$$

with  $R := 2\Pi - \mathbb{1}$  the reflection operator about the  $|0\rangle$  state of the ancilla and  $\Pi := |0\rangle\langle 0| \otimes \mathbb{1}$  the projector onto the ancilla  $|0\rangle$ , would have the effect [22]

$$\mathcal{Q}\mathcal{W}|0\rangle|\psi\rangle = |0\rangle U_{\mathbf{L}}|\psi\rangle. \quad (12)$$

Thus, we define

$$\mathcal{A} := \mathcal{Q}\mathcal{W} = -\mathcal{W}R\mathcal{W}^\dagger R\mathcal{W}. \quad (13)$$

We first discuss the requirement of  $s_{\mathbf{L}} = 2$ . Our form of the Taylor expansion leads to  $s_{\mathbf{L}}$  being of the form

$$s_{\mathbf{L}}(t) := \sum_{k=1}^{\infty} \frac{t^k}{k!} \prod_{j=1}^k \underbrace{\left[ \sum_{\ell_j=0}^{L_j-1} \alpha_{\ell_j} \right]}_{:=\Lambda_j} = \sum_{k=0}^{\infty} \frac{t^k}{k!} \prod_{j=1}^k \Lambda_j. \quad (14)$$

The restriction  $s_{\mathbf{L}} = 2$  therefore forces the simulation time  $t$  to be the only real root of

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} \prod_{j=1}^k \Lambda_j - 2 = 0 \quad (15)$$

which we call  $t_{\mathbf{L}}$ . If we were to include all orders in full, i.e.  $L_k = L$ ,  $\forall k$ , all  $\Lambda_j$  would be equal and the infinite sum on the left would become the series of the

exponential function. We call the time step for this case  $t_{\infty} = \log(2)/\Lambda$ , with the definition  $\Lambda := \sum_{j=0}^L \alpha_j$ .

Shorter times can be accomplished by using an extra qubit, as described in [22]. Since the only requirement for oblivious amplitude amplification to work is  $s_{\mathbf{L}} = 2$ , and shorter times mean  $s_{\mathbf{L}} < 2$ , i.e. the amplitude of the ancilla  $|0\rangle$  is too large, we can introduce an additional qubit to the ancilla and prepare it with enough weight such that the ancilla  $|0\rangle$  reduces to amplitude  $1/2$ . These shorter times are only relevant in the last time step of a simulation and have almost the same cost as a full step, so we limit the rest of the discussion to multiples of  $t_{\mathbf{L}}$ .

Equation (12) only strictly holds for unitary  $U_{\mathbf{L}}$ , but the series truncation means that  $U_{\mathbf{L}}$  is only approximately unitary. Therefore, we need the action of  $\mathcal{A}$  for a general  $U_{\mathbf{L}}$  and again follow [29]. Applying  $\mathcal{A}$  to a state  $|0\rangle|\psi\rangle$  and projecting onto the ancilla  $|0\rangle$  yields

$$\Pi\mathcal{A}|0\rangle|\psi\rangle = |0\rangle \left( \frac{3}{s_{\mathbf{L}}}U_{\mathbf{L}} - \frac{4}{s_{\mathbf{L}}^3}U_{\mathbf{L}}U_{\mathbf{L}}^\dagger U_{\mathbf{L}} \right) |\psi\rangle, \quad (16)$$

(derivation in Appendix B, Lemma 2) and we call the operator we are actually applying in the  $|\psi\rangle$  subspace

$$\tilde{\mathcal{A}}_{\mathbf{L}} := \frac{3}{s_{\mathbf{L}}}U_{\mathbf{L}} - \frac{4}{s_{\mathbf{L}}^3}U_{\mathbf{L}}U_{\mathbf{L}}^\dagger U_{\mathbf{L}}. \quad (17)$$

## 2.3 Gate construction

We also want to elaborate on the specific gate construction to implement  $\mathcal{A}$  efficiently, adapted from [29]. First, the ancilla is divided into  $\kappa + 1$  registers, where  $\kappa := \|\mathbf{L}\|_0$  is the number of non-zero elements in the vector  $\mathbf{L}$ . The first register is named  $q$  and contains  $\kappa$  qubits, while the others are given labels  $c_1 \dots c_{\kappa}$ , with  $c_k$  containing  $\lceil \log_2 L_k \rceil$  qubits.

The  $q$  register's purpose is to represent different orders, while registers  $c_k$  are needed for the terms in each order. This makes it convenient to use a multi-index  $j \equiv (k, \ell_1, \dots, \ell_k)$ . The corresponding state of the ancilla is

$$|j\rangle \equiv |k\rangle_q |\ell_1\rangle_{c_1} \dots |\ell_k\rangle_{c_k} \dots \quad (18)$$

where we leave the state of the registers  $c_{k'}$  with  $k' > k$  unspecified. The coefficient associated with this index is

$$\beta_j = \beta_{(k, \ell_1, \dots, \ell_k)} = \frac{t^k}{k!} \alpha_{\ell_1} \dots \alpha_{\ell_k}. \quad (19)$$

**PREPARE** For this operator, we slightly deviate from [29]. Exact implementation of  $\mathcal{P}$  as defined in Eq. (6) would necessitate the preparation of the  $c_k$  registers to be conditioned on qubits in the  $q$  register. We can, however, implement an operator  $\mathcal{P}^*$ , which acts equivalently to  $\mathcal{P}$  when used in  $\mathcal{W}$ , but is performed independently on each of the  $c_k$  registers without controls on the qubits in  $q$ .

The  $q$  register will contain the prefactor for each order  $k$  and uses unary coding, i.e.  $|k\rangle_q := |1^k 0^{k-\kappa}\rangle_q$ . Thus, the PREPARE operator  $\mathcal{P}^*(t)$  acts on this register proportional to

$$|0^\kappa\rangle_q \mapsto \sum_{k=0}^{\kappa} \sqrt{\frac{t^k}{k!} \prod_{j=1}^k \Lambda_j} |k\rangle_q. \quad (20)$$

This can be implemented by a rotation on the first qubit, and rotations controlled by the previous one on each subsequent qubit.

The  $c_k$  registers can now all be almost identically prepared to contain the coefficients of the Hamiltonian, where each index  $\ell$  is mapped to the qubits of  $c_k$  in regular binary coding. So the action of  $\mathcal{P}^*$  on a single register  $c_k$  is proportional to

$$|0\rangle_{c_k} \mapsto \sum_{\ell=0}^{L_k-1} \sqrt{\alpha_\ell} |\ell\rangle_{c_k}. \quad (21)$$

For this, any efficient method for arbitrary state preparation can be used, whose cost we discuss presently.

Combining these constituents into a single unitary  $\mathcal{P}^*$  and applying it to the whole ancilla yields the desired operator equivalent to Eq. (6) if used in  $\mathcal{W}$ , which is shown in more detail in Lemma 3 in Appendix B.

**SELECT** Using the established structure of the ancilla, the  $\mathcal{S}$  operator must have the action

$$\begin{aligned} \mathcal{S} |k\rangle_q |\ell_1\rangle_{c_1} \dots |\ell_k\rangle_{c_k} \dots |\ell_\kappa\rangle_{c_\kappa} |\psi\rangle \\ = |k\rangle_q |\ell_1\rangle_{c_1} \dots |\ell_k\rangle_{c_k} \dots |\ell_\kappa\rangle_{c_\kappa} \tilde{h}_{\ell_1} \dots \tilde{h}_{\ell_k} |\psi\rangle \end{aligned} \quad (22)$$

with  $\tilde{h}_\ell := -ih_\ell$ . This can be accomplished by having a sequence of groups of unitaries in the circuit.<sup>3</sup> Each of the groups  $m = 1 \dots \kappa$  contains the unitaries  $\tilde{h}_{\ell_m}$ , with  $\ell_m = 0 \dots L_m - 1$ , acting on the target state  $|\psi\rangle$ .

The register  $c_m$  is used as the addressing register for group  $m$ , i.e. the state  $|\ell_m\rangle_{c_m}$  determines which unitary in group  $m$  is applied. To achieve this, we use the fact that the  $c$  registers are in binary coding, so  $\ell_m$  is represented as a binary number with the  $\lceil \log_2 L_m \rceil$  qubits in  $c_m$  as digits. By controlling  $\tilde{h}_{\ell_m}$  on the  $c_m$  register in a way that matches the binary representation of  $\ell_m$ , only the unitary with the correct index is applied. For example,  $\tilde{h}_5$  would be controlled by the last and antepenultimate qubit in  $c_m$  and anti-controlled by all other qubits in  $c_m$  (since 5 corresponds to the state  $|0 \dots 0101\rangle$  in binary coding).

Additionally, the  $q$  register specifies how many of the groups are applied. If  $q$  is in the state  $|k\rangle_q$ , only the first  $k$  groups should be active. The unary coding in  $q$  makes this straightforward to implement by additionally controlling every unitary in group  $m$  with the  $m^{\text{th}}$  qubit in  $q$ . Figure 1 shows a sketch of the full construction.

<sup>3</sup>The groups in the circuit are numbered right-to-left to match the established numbering convention of the operators.

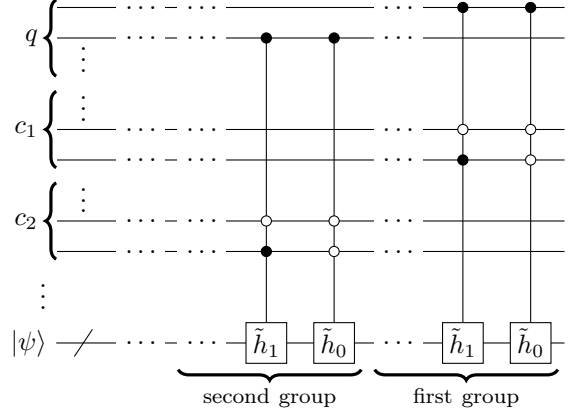


Figure 1: Sketch of the gate construction for  $\mathcal{S}$ . By taking advantage of the unary iteration structure, the  $T$ -count of the multi-controls can be significantly reduced [31]. However, we include this non-optimized diagram for pedagogical purposes.

**Gate cost** Lastly, we want to estimate the gate complexity of the operator  $\mathcal{A}$ . Its constituents are two reflections  $R$ , and three instances of  $\mathcal{W}$ , each of which contains one  $\mathcal{S}$  and two  $\mathcal{P}^*$ . For calculations using full orders as in [29], our analysis translates exactly to the gate construction given there. We consider the universal set of Clifford +  $T$  and count the number of expensive  $T$ -gates [35–37] for our complexity analysis.

Each reflection  $R$  is a single Pauli- $Z$  operator on one of the ancilla qubits (padded between two NOT gates), anti-controlled on all others. This can be done with  $\mathcal{O}(\sum_k \log_2 L_k)$   $T$ -gates and a second ancilla register of size  $(\kappa + \sum_k \lceil \log_2 L_k \rceil - 2)$  [38].

The PREPARE stage for the  $q$  register consists of  $\kappa - 1$  controlled rotations with a total  $T$ -complexity of  $\mathcal{O}(\kappa)$ . Each of the  $\kappa$  registers  $c_k$  needs to be initialized to a specific state with  $2^{\lceil \log_2 L_k \rceil} \sim L_k$  coefficients, requiring between  $\mathcal{O}(\sum_k L_k)$  and  $\mathcal{O}(\sum_k \sqrt{L_k} \log^2(L_k/\epsilon))$   $T$ -gates per register, depending on the number of additionally available ancillas, where  $\epsilon$  is the accuracy of the preparation [34]. In total, this yields a  $T$ -count between  $\mathcal{O}(\sum_k L_k)$  and  $\mathcal{O}(\sum_k \sqrt{L_k} \log^2(L_k/\epsilon))$ .

The fact that the controls of each  $h_\ell$  in  $\mathcal{S}$  form a so-called unary iteration can be exploited to lower the  $T$ -gate count. Each sequence of  $L_k$  operators can be implemented using  $\mathcal{O}(L_k)$   $T$ -gates [31], plus  $L_k$  times the cost of performing a single  $-ih_\ell$  operator, totalling to  $\sum_k L_k$  such operators. Thus, the  $T$ -complexity of  $\mathcal{S}$  for generic Hamiltonians in the form of Eq. (1) is of  $\mathcal{O}(\sum_k L_k)$ , which we will use in this paper. Moreover, recent work [28] has shown that a SELECT process can be yet more efficient for the special case of  $N$ -orbital electronic structure problems, where the  $T$ -complexity is as low as  $\mathcal{O}(nN)$ , with  $n = \max\{k : L_k \neq 0\}$ .

Combining all these counts results in a total complexity of  $\mathcal{O}(\sum_k L_k)$  for  $\mathcal{A}$ . As a proxy to use for the total gate cost in our results we thus define

$$C_{\mathbf{L}} := \sum_{k=1}^{\infty} L_k = \|\mathbf{L}\|_1. \quad (23)$$

This definition includes the cost of a full expansion to order  $n$  as the special case  $C_n = nL$ , consistent with our previous notation. From this cost of a single time step, we discuss the complexity  $C_\epsilon$  to reach some desired total simulation error  $\epsilon$  in the next subsection.

## 2.4 Error bounds

We consider the error of the method per time step to be the norm of an operator  $\Delta_{\mathbf{L}}$  which fulfills

$$U(t_\infty) = \tilde{\mathcal{A}}_{\mathbf{L}}(t_\infty) + \Delta_{\mathbf{L}}(t_\infty) \quad (24)$$

where we now use the step size  $t_\infty = \log(2)/\Lambda$ . We find that the error made by applying  $\Pi\mathcal{A}$  once and tracing out the ancilla can be bounded by<sup>4</sup>

$$\delta_{\mathbf{L}} := \|\Delta_{\mathbf{L}}(t_\infty)\| \leq 2 - s_{\mathbf{L}}(t_\infty) =: \varepsilon_{\mathbf{L}} \quad (25)$$

up to order  $\varepsilon_{\mathbf{L}}$  (details in Appendix B, Lemma 4). Because using  $t_{\mathbf{L}}$  or  $t_\infty$  makes no difference in the error up to order  $\varepsilon_{\mathbf{L}}$ , we exclusively use  $t_\infty$  in our calculations. The error for a total simulation time  $\tau = rt_\infty = r \log(2)/\Lambda$ ,  $r \in \mathbb{N}$ , is then

$$\|\tilde{\mathcal{A}}_{\mathbf{L}}(t_\infty)^r - U(t_\infty)^r\| = r\delta_{\mathbf{L}} = \frac{\Lambda\delta_{\mathbf{L}}}{\log 2} \tau \leq r\varepsilon_{\mathbf{L}}, \quad (26)$$

also up to order  $\varepsilon_{\mathbf{L}}$  (see Appendix B, Lemma 5).

We call the bound on the total simulation error of  $r$  steps  $\epsilon := r\varepsilon$ . The  $T$ -gate complexity  $C_\epsilon$  of a simulation for time  $\tau$  in terms of the total error bound  $\epsilon$  is then in the range

$$\mathcal{O}\left(\frac{\Lambda\tau \log \frac{\Lambda\tau}{\epsilon}}{\log \log \frac{\Lambda\tau}{\epsilon}}\right) < C_\epsilon \leq \mathcal{O}\left(\frac{L\Lambda\tau \log \frac{\Lambda\tau}{\epsilon}}{\log \log \frac{\Lambda\tau}{\epsilon}}\right), \quad (27)$$

depending on the Hamiltonian. This is shown in detail in Appendix B, Corollary 10, which makes use of Lemmas 6 and 8 and Corollaries 7 and 9.

## 2.5 Insertion strategy

The notion of partially included orders together with an expression for the error bound allows us to start from any given expansion  $\mathbf{L}$  and determine which  $L_k$  should be increased by 1 – i.e. which additional gate should be included – to give the quickest decrease of the error bound. Specifically, it is the  $k$  which maximizes the expression

$$\sum_{\nu \geq k} \frac{t^\nu}{\nu!} \alpha_{1+L_k} \prod_{\substack{j \neq k \\ 1 \leq j \leq \nu}} \left( \sum_{i=1}^{L_j} \alpha_i \right). \quad (28)$$

Starting from  $\mathbf{L} = \mathbf{0}$ , repeatedly adding terms that maximize (28) results in a greedy algorithm for decreasing the error bound, which we used to iteratively construct circuits.

<sup>4</sup> $\|\cdot\|$  in this paper always means the operator norm.

## 3 Results

We first observe that for Hamiltonians with evenly distributed magnitudes  $\alpha_\ell$ , the only benefit of our modification is the finer control over the total gate count. By construction, whenever  $C_{\mathbf{L}} = \nu L$ ,  $\nu \in \mathbb{N}$ , our protocol and the method used in [29] yield identical results.

We may expect our modification to be advantageous whenever the magnitudes of  $\alpha_\ell$  vary over several orders of magnitude because this allows terms in low orders containing small  $\alpha_\ell$  to be smaller than terms in higher orders containing large  $\alpha_\ell$ . Such magnitude distributions are often found in electronic structure

Table 1: Molecules used in our calculations with their molecular formula, PubChem Compound ID (CID), number of qubits (excluding ancillas), and number of terms  $L$ .

Formula	CID	Qubits	$L$
HO	157 350	12	631
HF	16 211 014	12	631
HN	5 460 607	12	631
LiH	62 714	12	631
BH	6 397 184	12	631
BeH <sub>2</sub>	139 073	14	666
CH <sub>2</sub>	123 164	14	1086
NH <sub>2</sub>	123 329	14	1086
BH <sub>2</sub>	139 760	14	1086
H <sub>2</sub> O	962	14	1086
BH <sub>3</sub>	6331	16	1953
CH <sub>3</sub>	3 034 819	16	1969
NH <sub>3</sub>	222	16	2929
CH <sub>4</sub>	297	18	6892
O <sub>2</sub>	977	20	2239
N <sub>2</sub>	947	20	2951
NO	145 068	20	4427
CN	5 359 238	20	5835
BeO	14 775	20	5851
LiF	224 478	20	5851
CO	281	20	5851
BN	66 227	20	5851
LiOH	3939	22	8734
HBO	518 615	22	8758
HCN	768	22	8758
HOF	123 334	22	12 070
CHO	123 370	22	12 070
CHF	186 213	22	12 074
HNO	945	22	12 078
H <sub>2</sub> NO	5 460 582	24	9257
CH <sub>2</sub> O	712	24	9257
NH <sub>2</sub> F	139 987	24	15 673
CH <sub>2</sub> F	138 041	24	15 681
CH <sub>3</sub> F	11 638	26	18 600
CH <sub>3</sub> Li	2 724 049	26	19 548
H <sub>3</sub> NO	787	26	22 080
OCH <sub>3</sub>	123 146	26	39 392
LiBH <sub>4</sub>	4 148 881	28	27 473
CH <sub>3</sub> OH	887	28	30 419
C <sub>4</sub> H <sub>8</sub> O <sub>2</sub>	8857	76	1 614 647
C <sub>8</sub> H <sub>6</sub>	12 302 244	92	1 897 809

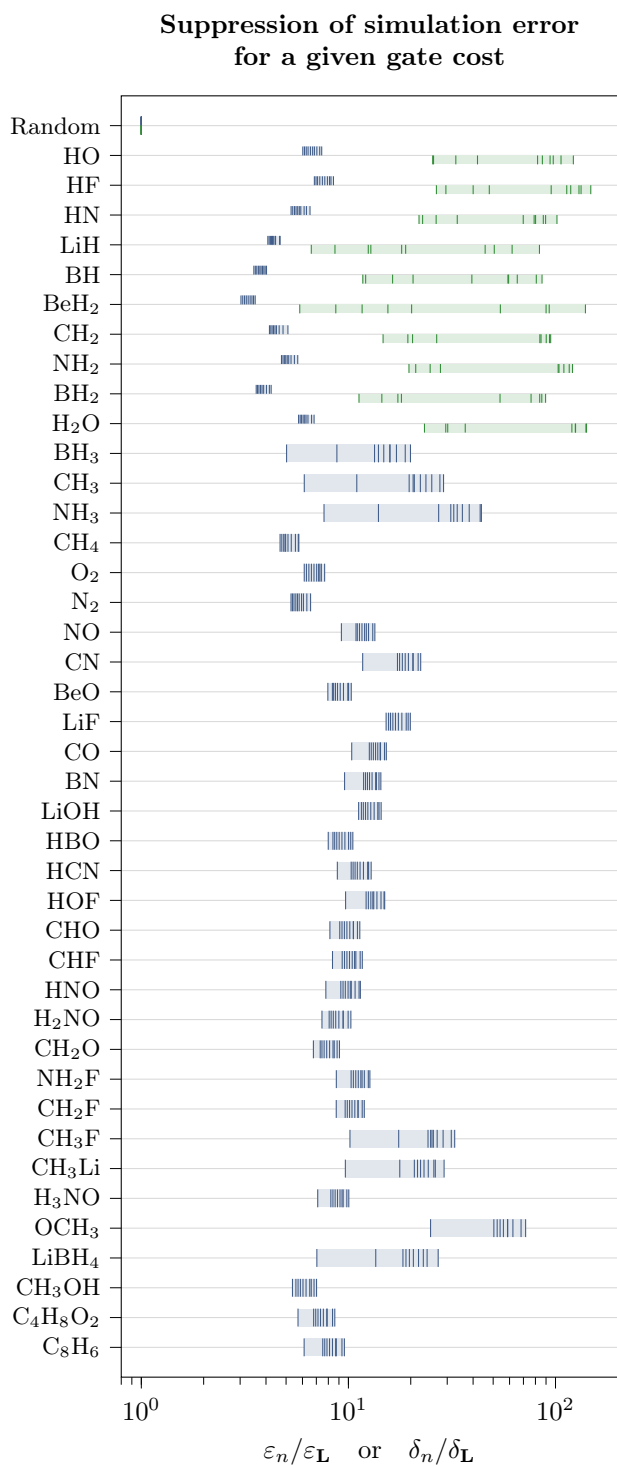


Figure 2: Ratio of the errors obtained without, versus with, our modification for different molecules at identical implementation cost, using a time step of  $t_\infty$ . Errors were evaluated at cost values  $C_L = (1 \dots 10)L$ . Each vertical line represents the ratio of errors at some cost  $C_L$ , the shaded areas indicate the range from the smallest to the largest data point. The advantage of our modified algorithm therefore increases left-to-right. Top-bottom split data indicates ratios of error bounds  $\varepsilon_n/\varepsilon_L$  at the top (marked in blue) and ratios of numerically obtained errors  $\delta_n/\delta_L$  at the bottom (marked in green). If there is no split, the blue lines represent error bounds only.

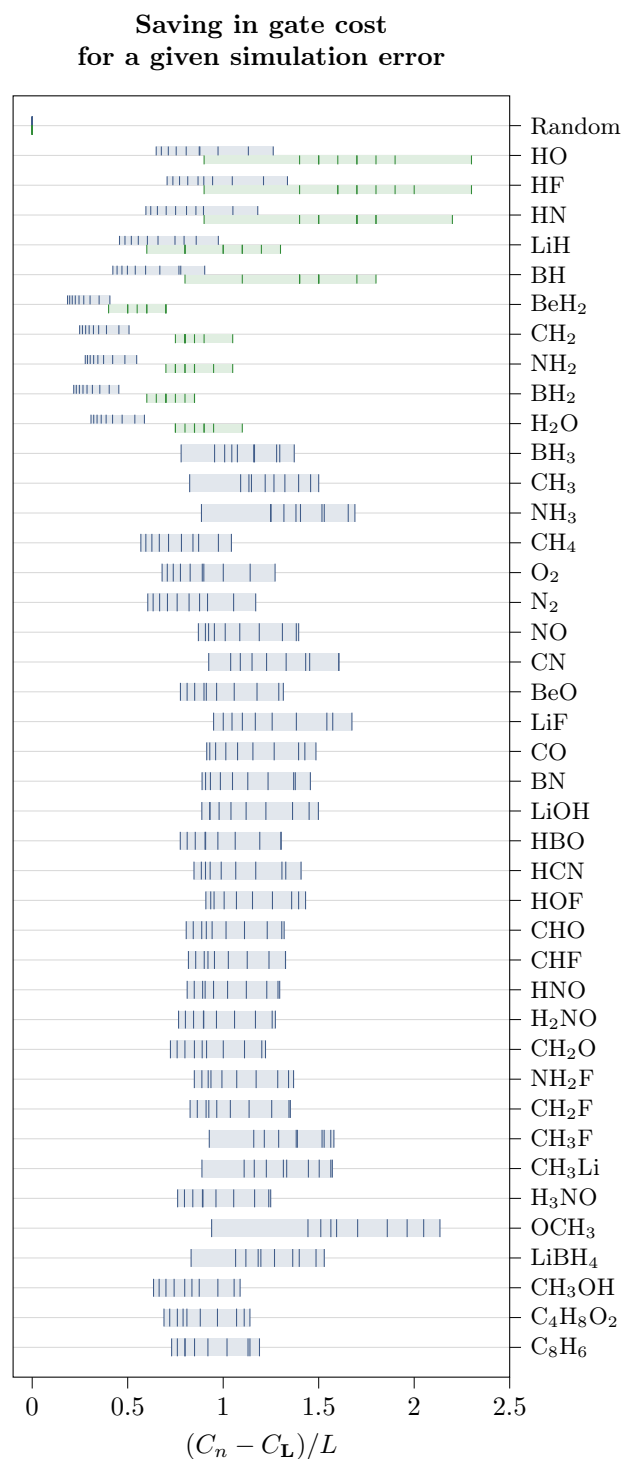


Figure 3: Difference between the cost of full expansions  $C_n = nL$  to order  $n = 1 \dots 10$  and the cost of an iteratively constructed circuit  $C_L$  to arrive at the same error, normalized to the cost of one full order  $L$ , for each molecule. Each vertical line represents the difference at some value of  $n$ , the shaded areas indicate the range from the smallest to the largest data point. The advantage of our modified algorithm therefore increases left-to-right. The time step size is  $t_\infty$ . Top-bottom split data indicates differences for error bounds at the top (marked in blue) and for numerically obtained errors on the bottom (marked in green). If there is no split, the blue lines represent error bounds only.

Hamiltonians for molecules [39]. Because the efficiency of our method depends critically on the specific amplitudes in the Hamiltonian, analytical results are hard to obtain. Therefore we resort to a numerical study comparing the accuracy of the modification to the method in [29] for a group of molecules listed in Table 1.

The Hamiltonians for these molecules were obtained using OpenFermion [40] and PySCF [41], with the basis set STO-3G [42], and geometry data retrieved from PubChem [43] and the NIST Computational Chemistry Comparison and Benchmark Database [44]. Mapping from second quantization to spin operators was done using the Jordan-Wigner transformation [45].

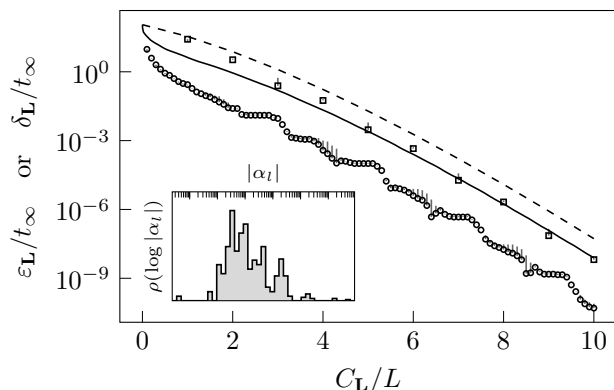


Figure 4: Accuracy of the Taylor expansion for the electronic Hamiltonian of hydrogen fluoride (HF), at time step size  $t_\infty$ , in terms of the error per unit time vs the circuit cost  $C_L$  as defined in Eq. (23) per cost of a full order. Lines are the error bounds  $\varepsilon_L$  for the unmodified --- and modified — circuit. Squares  $\square$  are the numerically obtained errors  $\delta_L$  for fully expanded orders, circles  $\circ$  analogous for partial orders. The vertical gray bars point to where the error would be if we could implement  $U_L$  without the amplification step. The inset shows the distribution  $\rho$  of the logarithms of weights in the Hamiltonian  $\log|\alpha_\ell|$ .

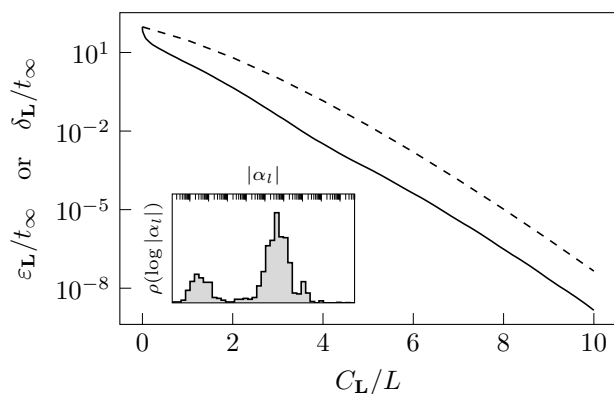


Figure 5: Identical plot to Fig. 4 showing only the convergence of error bounds for ammonia ( $\text{NH}_3$ ). Notice the increasing distance between the unmodified --- and modified — variants of the algorithm between the orders 2 and 4, which is not present in Fig. 4. It is caused by the two distinct clusters in the distribution of the logarithmic weights in the Hamiltonian visible in the inset.

To showcase that our method yields improvements regardless of the basis set used, we also performed calculations using the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets [46, 47] for  $\text{H}_2$  and  $\text{LiH}$ . The results are shown in Appendix A.

In addition to the listed molecules, we also replaced the coefficients of the Hamiltonian for  $\text{LiH}$  with random numbers from a normal distribution with  $\mu = 1$  and  $\sigma = 0.1$ , to show the vanishing effect of our modification whenever all weights are similar. These results are labeled *Random*.

Figure 4 shows the error bounds as well as the numerically evaluated exact errors per unit time for hydrogen fluoride. Compared to the expansion to full orders, we see that our modification leads to a much quicker decrease of the error bound as well as the exact error in the range  $0 < C_L < L$ , followed by very similar convergence for  $C_L > L$ . This pattern is consistent with the convergence we observed for most other molecules we calculated.

Some compounds in our set – namely  $\text{BH}_3$ ,  $\text{CH}_3$ ,  $\text{NH}_3$ ,  $\text{CH}_3\text{F}$ ,  $\text{CH}_3\text{Li}$ ,  $\text{OCH}_3$ , and  $\text{LiBH}_4$  – show a slightly different behavior, where the ratio of error bounds using the modified and unmodified versions increases once more later in the iteration. Figure 5 illustrates this using  $\text{NH}_3$  as an example. The delayed convergence is caused by a distinct second peak in the distribution of the logarithms of weights in the Hamiltonian  $\log|\alpha_\ell|$ , present in the mentioned molecules. The rest of our set shows distributions rather similar to that of HF in the inset in Fig. 4. This disparity is also visible in the summarized results in Figs. 2 and 3, where the spread for the mentioned molecules (especially for the error in Fig. 2) is much greater than for the rest.

To summarize the results for all molecules, we obtained the ratio of the errors of the original and the modified version at cost values  $C_L = nL$ , with  $n = 1 \dots 10$ , where the time step was set to  $t_\infty$  for each respective molecule. The results are depicted in Fig. 2. Across the listed molecules, our modification consistently yields errors roughly one order of magnitude lower than the unmodified method at equivalent costs, with some ratios as low as 3 and some as high as 100.

We also compared the cost required to obtain a certain error threshold. To this end, the errors  $\delta_n$  of the expansions to full orders  $n$  were calculated, and the cost  $C_L$  of the modified version to yield the same error was recorded. The results are depicted in Fig. 3. Using the modified method leads to saving approximately one order in most cases, i.e. the accuracy obtained by expanding  $n$  full orders can be produced with a cost of  $C_L = (n - 1)L$ .

Our results show no strong correlation with neither the number of qubits in the Hamiltonian nor the number of terms  $L$ . Therefore we presume that these properties will also hold for other chemical Hamiltonians obtained in the same way.



## 4 Conclusion

We demonstrated that a natural extension of the method proposed in [29] can lead to noticeable improvements in the convergence of the approximation when used for electronic structure Hamiltonians of molecules. The asymptotic behavior is equivalent; however, minimizing the required number of gates will be important for implementations on actual quantum hardware.

Our modification does not need the introduction of any new subroutines. It only rearranges the gate construction to facilitate quicker convergence of both the error bound and the actual error.

Due to the lack of analytic relations between the amplitudes in the Hamiltonians we investigated, only numeric results are available. However, because of the relatively large sample size of molecules we considered, it stands to reason that this behavior will generalize to a large portion of electronic structure Hamiltonians.

As mentioned in the introduction, in light of other methods like qubitization and quantum signal processing, the truncated Taylor series may be of particular relevance for diagonally dominant and time-dependent Hamiltonians. Investigating the suitability of our modification to such problems would therefore be an interesting question for future work.

Furthermore, combining our proposed adaptations with the improvements by Novo and Berry [48], who add an additional correction step to the method, could also be worth exploring.

## Acknowledgments

The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility (DOI 10.5281/zenodo.22558) in carrying out this work. ETC was supported by the EPSRC (grant no. EP/M024261/1). SCB acknowledges support from the EU Flagship project AQ-TION, the NQIT Hub (EP/M013243/1), and the QCS Hub (EP/T001062/1). The authors also thank Bálint Koczor and Sam McArdle for useful discussions, as well as Yingkai Ouyang for helpful comments on the manuscript.

## References

- [1] R. P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, Jun 1982. DOI 10.1007/BF02650179.
- [2] S. Lloyd. Universal quantum simulators. *Science*, 273:1073–1078, Aug 1996. DOI 10.1126/science.273.5278.1073.
- [3] D. S. Abrams and S. Lloyd. Simulation of many-body Fermi systems on a universal quantum computer. *Physical Review Letters*, 79(13):2586–2589, Sep 1997. DOI 10.1103/PhysRevLett.79.2586.
- [4] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme. Quantum algorithms for fermionic simulations. *Physical Review A*, 64(2), Jul 2001. DOI 10.1103/PhysRevA.64.022319.
- [5] A. Aspuru-Guzik. Simulated quantum computation of molecular energies. *Science*, 309(5741):1704–1707, Sep 2005. DOI 10.1126/science.1113479.
- [6] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders. Efficient quantum algorithms for simulating sparse Hamiltonians. *Communications in Mathematical Physics*, 270(2):359–371, Dec 2006. DOI 10.1007/s00220-006-0150-x.
- [7] H. Wang, S. Kais, A. Aspuru-Guzik, and M. R. Hoffmann. Quantum algorithm for obtaining the energy spectrum of molecular systems. *Physical Chemistry Chemical Physics*, 10(35):5388, 2008. DOI 10.1039/b804804e.
- [8] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik. Simulation of electronic structure Hamiltonians using quantum computers. *Molecular Physics*, 109(5):735–750, 2011. DOI 10.1080/00268976.2011.552441.
- [9] E. Campbell. Random compiler for fast Hamiltonian simulation. *Physical Review Letters*, 123(7):070503, Aug 2019. DOI 10.1103/physrevlett.123.070503.
- [10] A. M. Childs and Y. Su. Nearly optimal lattice simulation by product formulas. *Physical Review Letters*, 123(5):050503, Aug 2019. DOI 10.1103/physrevlett.123.050503.
- [11] A. M. Childs, A. Ostrander, and Y. Su. Faster quantum simulation by randomization. *Quantum*, 3:182, Sep 2019. DOI 10.22331/q-2019-09-02-182.
- [12] Y. Ouyang, D. R. White, and E. T. Campbell. Compilation by stochastic Hamiltonian sparsification. *Quantum*, 4:235, Feb 2020. DOI 10.22331/q-2020-02-27-235.
- [13] H. F. Trotter. On the product of semi-groups of operators. *Proceedings of the American Mathematical Society*, 10(4):545–551, 1959. DOI 10.2307/2033649.
- [14] M. Suzuki. Generalized Trotter’s formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Communications in Mathematical Physics*, 51(2):183–190, 1976. DOI 10.1007/BF01609348.
- [15] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer. Gate-count estimates for performing quantum chemistry on small quantum computers. *Physical Review A*, 90:022305, Aug 2014. DOI 10.1103/PhysRevA.90.022305.
- [16] R. Babbush, J. McClean, D. Wecker, A. Aspuru-Guzik, and N. Wiebe. Chemical basis of Trotter-Suzuki errors in quantum chemistry simulation. *Physical Review A*, 91(2), Feb 2015. DOI 10.1103/PhysRevA.91.022311.

- [17] M. B. Hastings, D. Wecker, B. Bauer, and M. Troyer. Improving quantum algorithms for quantum chemistry, 2014. arXiv 1403.1539.
- [18] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doherty, and M. Troyer. The Trotter step size required for accurate quantum simulation of quantum chemistry, 2014. arXiv 1406.4920.
- [19] J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik. Exploiting locality in quantum computation for quantum chemistry. *The Journal of Physical Chemistry Letters*, 5(24):4368–4380, Dec 2014. DOI 10.1021/jz501649m.
- [20] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu. Theory of Trotter Error with Commutator Scaling. *Physical Review X*, 11:011020, Feb 2021. DOI 10.1103/PhysRevX.11.011020.
- [21] A. M. Childs and N. Wiebe. Hamiltonian simulation using linear combinations of unitary operations. *Quantum Information & Computation*, 12(11–12):901–924, Nov 2012. DOI 10.26421/QIC12.11-12.
- [22] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma. Exponential improvement in precision for simulating sparse Hamiltonians. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, page 283–292, New York, NY, USA, 2014. Association for Computing Machinery. DOI 10.1145/2591796.2591854.
- [23] G. H. Low, V. Kliuchnikov, and N. Wiebe. Well-conditioned multiproduct Hamiltonian simulation, 2019. arXiv 1907.11679.
- [24] D. W. Berry, A. M. Childs, and R. Kothari. Hamiltonian Simulation with nearly optimal dependence on all parameters. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 792–809, 2015. DOI 10.1109/FOCS.2015.54.
- [25] G. H. Low and I. L. Chuang. Optimal Hamiltonian simulation by quantum signal processing. *Physical Review Letters*, 118(1):010501, Jan 2017. DOI 10.1103/physrevlett.118.010501.
- [26] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC '19, pages 193–204. ACM Press, 2019. DOI 10.1145/3313276.3316366.
- [27] G. H. Low and I. L. Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, Jul 2019. DOI 10.22331/q-2019-07-12-163.
- [28] D. W. Berry, C. Gidney, M. Motta, J. R. McClean, and R. Babbush. Qubitization of arbitrary basis quantum chemistry leveraging sparsity and low rank factorization. *Quantum*, 3:208, Dec 2019. DOI 10.22331/q-2019-12-02-208.
- [29] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma. Simulating Hamiltonian dynamics with a truncated Taylor series. *Physical Review Letters*, 114:090502, Mar 2015. DOI 10.1103/PhysRevLett.114.090502.
- [30] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su. Toward the first quantum simulation with quantum speedup. *Proceedings of the National Academy of Sciences*, 115(38):9456–9461, Sep 2018. DOI 10.1073/pnas.1801723115.
- [31] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven. Encoding electronic spectra in quantum circuits with linear T complexity. *Physical Review X*, 8:041015, Oct 2018. DOI 10.1103/PhysRevX.8.041015.
- [32] G. H. Low and N. Wiebe. Hamiltonian simulation in the interaction picture, 2018. arXiv 1805.00675.
- [33] R. Babbush, D. W. Berry, J. R. McClean, and H. Neven. Quantum simulation of chemistry with sublinear scaling in basis size. *npj Quantum Information*, 5(1), Nov 2019. DOI 10.1038/s41534-019-0199-y.
- [34] G. H. Low, V. Kliuchnikov, and L. Schaeffer. Trading T-gates for dirty qubits in state preparation and unitary synthesis, 2018. arXiv 1812.00954.
- [35] S. Bravyi and A. Kitaev. Universal quantum computation with ideal Clifford gates and noisy ancillas. *Phys. Rev. A*, 71:022316, Feb 2005. DOI 10.1103/PhysRevA.71.022316.
- [36] M. Howard, J. Wallman, V. Veitch, and J. Emerson. Contextuality supplies the magic for quantum computation. *Nature*, 510(7505):351–355, Jun 2014. DOI 10.1038/nature13460.
- [37] E. T. Campbell, B. M. Terhal, and C. Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, Sep 2017. DOI 10.1038/nature23460.
- [38] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*, chapter 4, pages 182–184. Cambridge University Press, 2nd edition, 2010. ISBN 978 1 107 00217 3.
- [39] T. Helgaker, P. Jorgensen, and J. Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2000. ISBN 978 1 118 53147 1.
- [40] J. R. McClean, K. J. Sung, I. D. Kivlichan, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, O. Higgott, C. Huang, J. Izaac, Z. Jiang, X. Liu, S. McArdle, M. Neeley, T. O’Brien, B. O’Gorman, I. Ozfidan, M. D. Radin, J. Romero, N. Rubin, N. P. D. Sawaya, K. Setia, S. Sim, D. S. Steiger, M. Steudtner, Q. Sun, W. Sun, D. Wang, F. Zhang, and R. Babbush. OpenFermion: The electronic structure package for quantum computers, 2017. arXiv 1710.07629.

- [41] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K. Chan. PySCF: the Python-based simulations of chemistry framework, 2017. DOI 10.1002/wcms.1340.
- [42] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of Slater-type atomic orbitals. *The Journal of Chemical Physics*, 51(6): 2657–2664, 1969. DOI 10.1063/1.1672392.
- [43] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, and E. E. Bolton. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Research*, 47(D1):D1102–D1109, Oct 2018. DOI 10.1093/nar/gky1033.
- [44] R. D. Johnson III. NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101, Release 20, Aug 2019. URL <http://cccbdb.nist.gov>.
- [45] P. Jordan and E. Wigner. Über das Paulische Äquivalenzverbot. *Zeitschrift für Physik*, 47(9): 631–651, 1928. DOI 10.1007/BF01331938.
- [46] T. H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of chemical physics*, 90(2):1007–1023, 1989. DOI 10.1063/1.456153.
- [47] B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, and A. K. Wilson. Gaussian basis sets for use in correlated molecular calculations. VII. Valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg. *Theoretical Chemistry Accounts*, 128(1):69–82, 2011. DOI 10.1007/s00214-010-0764-0.
- [48] L. Novo and D. W. Berry. Improved Hamiltonian simulation via a truncated Taylor series and corrections, 2016. arXiv 1611.10033.

## A Effect of basis set choice

To ensure the effect we observed also holds for larger basis sets, we additionally performed calculations for H<sub>2</sub> and LiH with cc-pVDZ, cc-pVTZ, and cc-pVQZ sets.

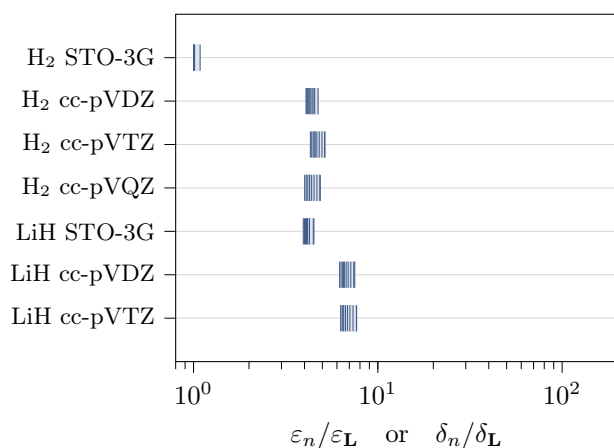


Figure 6: Ratio of the error bounds obtained without, versus with, our modification for different molecules and basis sets, using a time step of  $t_\infty$ . Errors were evaluated at each cost value  $C_L = (1 \dots 10)L$ . Each line represents the ratio of error bounds  $\varepsilon_n/\varepsilon_L$  at some cost  $C_L$ . The advantage of our modified algorithm therefore increases left-to-right.

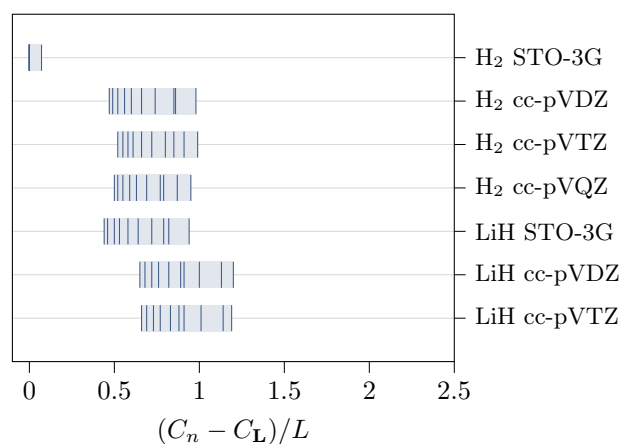


Figure 7: Difference of the cost of full expansions  $C_n = nL$  to order  $n = 1 \dots 10$  and the cost of an iteratively constructed circuit  $C_L$  to arrive at the same error bound, normalized to the cost of one full order  $L$ , for each molecule and basis set. The advantage of our modified algorithm therefore increases left-to-right. The time step size is  $t_\infty$ .

Figures 6 and 7 show that for the considered cases, larger basis sets seem to slightly enhance the advantage of our modification. As a special case, for H<sub>2</sub> our proposed method yields almost no improvement when using STO-3G, due to the very low number of only 15 terms. Apart from this outlier, the influence of the choice of basis set on our algorithm’s performance seems small.

## B Proofs

For completeness and convenience, we collect several of the results we used in this appendix, including some that may be well known.

**Lemma 1.** *The optimal choice for the number of amplification steps is  $\nu = 1$ , resulting in  $s_{\mathbf{L}} = 2$ .*

*Proof.* For unitary  $U_{\mathbf{L}}$ , the operator  $\mathcal{Q}^\nu \mathcal{W}$ , with  $\mathcal{Q}$  as defined in Eq. (11), would have the effect [22]

$$\begin{aligned} \mathcal{Q}^\nu \mathcal{W} |0\rangle |\psi\rangle &= \sin[(2\nu + 1) \sin^{-1}(s_{\mathbf{L}}^{-1})] |0\rangle U_{\mathbf{L}} |\psi\rangle \\ &\quad + \cos[(2\nu + 1) \cos^{-1}(N)] |0^\perp, \Phi\rangle. \end{aligned}$$

For any given number of amplification steps  $\nu$ , the amplitude of the desired state  $|0\rangle U_{\mathbf{L}} |\psi\rangle$  can be tuned to 1 by setting  $t$  such that  $s_{\mathbf{L}}$  fulfills

$$s_{\mathbf{L}} = \sin\left(\frac{\pi}{4\nu + 2}\right)^{-1} \sim \frac{4\nu + 2}{\pi}. \quad (29)$$

For this argument it is sufficient to analyze the full expansion to order  $n$ . To find the optimal number  $\nu$  we consider the operator  $\mathcal{Q}^\nu \mathcal{W}$ , which contains the most expensive operator  $\mathcal{W}$  a total of  $2\nu + 1$  times. Therefore the cost is approximately linear in  $\nu$ , meaning it is also linear in  $s_n$ . However, we know that

$$s_n = \sum_{k=0}^n \frac{t^k}{k!} \underbrace{\left(\sum_{\ell=0}^{L-1} \alpha_\ell\right)^k}_{:=\Lambda} = \sum_{k=0}^n \frac{t^k \Lambda^k}{k!} \approx e^{\Lambda t}, \quad (30)$$

where the rightmost approximation holds for sufficiently large  $n$ . Equations (29) and (30) imply that the cost is exponential in  $t$ , indicating there is no benefit in amplifying more than once. Exact numerical evaluation shows that for  $n \rightarrow \infty$ , one or two amplification steps ( $\nu \in \{1, 2\}$ ) yield approximately equivalent time-per-gate, but the smaller  $t$  of  $\nu = 1$  leads to quicker convergence in  $n$ . Consequently, it is best to choose  $\nu = 1$ . This choice forces  $s_{\mathbf{L}}$  to satisfy

$$s_{\mathbf{L}} = \sin\left(\frac{\pi}{6}\right)^{-1} = 2.$$

□

**Lemma 2.** *The action of  $\Pi\mathcal{A}$  on a product state  $|0\rangle |\psi\rangle$  is given by [29]*

$$\Pi\mathcal{A} |0\rangle |\psi\rangle = |0\rangle \left( -\frac{4}{s_{\mathbf{L}}^3} U_{\mathbf{L}} U_{\mathbf{L}}^\dagger U_{\mathbf{L}} + \frac{3}{s_{\mathbf{L}}} U_{\mathbf{L}} \right) |\psi\rangle.$$

*Proof.* We can explicitly expand  $R$  and use that  $\Pi^2 = \Pi$  as well as  $\Pi |0\rangle |\psi\rangle = |0\rangle |\psi\rangle$  to find

$$\begin{aligned} \Pi\mathcal{A} |0\rangle |\psi\rangle &= -\Pi\mathcal{W}R\mathcal{W}^\dagger R\mathcal{W} |0\rangle |\psi\rangle \\ &= -\Pi\mathcal{W}(2\Pi - \mathbf{1})\mathcal{W}^\dagger(2\Pi - \mathbf{1})\mathcal{W} |0\rangle |\psi\rangle \\ &= (-4\Pi\mathcal{W}\Pi\mathcal{W}^\dagger\Pi\mathcal{W} + 2\Pi\mathcal{W}\mathcal{W}^\dagger\Pi\mathcal{W} + 2\Pi\mathcal{W}\Pi\mathcal{W}^\dagger\mathcal{W} - \Pi\mathcal{W}\mathcal{W}^\dagger\mathcal{W}) |0\rangle |\psi\rangle \\ &= (-4\Pi\mathcal{W}\Pi\mathcal{W}^\dagger\Pi\mathcal{W} + 3\Pi\mathcal{W}) |0\rangle |\psi\rangle \\ &= (-4 \underbrace{\Pi\mathcal{W}\Pi\mathcal{W}^\dagger\Pi\mathcal{W}\Pi}_{\frac{1}{s_{\mathbf{L}}}(|0\rangle\langle 0| \otimes U_{\mathbf{L}})} + 3\Pi\mathcal{W}\Pi) |0\rangle |\psi\rangle \\ &= |0\rangle \left( -\frac{4}{s_{\mathbf{L}}^3} U_{\mathbf{L}} U_{\mathbf{L}}^\dagger U_{\mathbf{L}} + \frac{3}{s_{\mathbf{L}}} U_{\mathbf{L}} \right) |\psi\rangle \end{aligned}$$

as claimed. □

**Lemma 3.** If used in  $\mathcal{W}|0\rangle|\psi\rangle$ ,  $\mathcal{P}^*$  has the same effect as  $\mathcal{P}$ , i.e.  $\mathcal{W}|0\rangle|\psi\rangle = (\mathcal{P}^\dagger \otimes \mathbb{1}) \mathcal{S}(\mathcal{P} \otimes \mathbb{1})|0\rangle|\psi\rangle = (\mathcal{P}^{\star\dagger} \otimes \mathbb{1}) \mathcal{S}(\mathcal{P}^* \otimes \mathbb{1})|0\rangle|\psi\rangle$

*Proof.*  $\mathcal{P}^*$  on any of the  $c_k$  registers has the action

$$\mathcal{P}^*|0\rangle_{c_k} = \frac{1}{\sqrt{\Lambda_k}} \sum_{\ell=0}^{L_k-1} \sqrt{\alpha_\ell} |\ell\rangle_{c_k}$$

and on the  $q$  register

$$\mathcal{P}^*|0\rangle_q = \frac{1}{\sqrt{N_q}} \sum_{k=0}^{\kappa} \sqrt{\frac{t^k}{k!} \prod_{j=1}^k \Lambda_j} |k\rangle_q$$

where  $N_q = \sum_{k=0}^{\infty} \frac{t^k}{k!} \prod_{j=1}^k \Lambda_j$  and  $\kappa$  is the largest nonzero index in  $\mathbf{L}$ . Therefore

$$\mathcal{S}(\mathcal{P}^* \otimes \mathbb{1})|0\rangle|\psi\rangle = \frac{1}{\sqrt{N_q} \prod_{k=1}^{\kappa} \Lambda_k} \sum_{k=0}^{\kappa} \sqrt{\left( \prod_{j=1}^k \Lambda_j \right) \frac{t^k}{k!} |k\rangle_q} \bigotimes_{j=1}^k \left( \sum_{\ell=0}^{L_j-1} |\ell\rangle_{c_j} \sqrt{\alpha_\ell \tilde{h}_\ell} \right) \bigotimes_{j=k+1}^{\kappa} \left( \sum_{\ell=0}^{L_j-1} |\ell\rangle_{c_j} \sqrt{\alpha_\ell} \right) |\psi\rangle$$

Transforming back with  $\mathcal{P}^{\star\dagger}$  and projecting onto the ancilla  $|0\rangle$  yields

$$\begin{aligned} \Pi(\mathcal{P}^{\star\dagger} \otimes \mathbb{1}) \mathcal{S}(\mathcal{P}^* \otimes \mathbb{1}) &= \frac{\Pi}{N_q \prod_{k=1}^{\kappa} \Lambda_k} \sum_{k=0}^{\kappa} \left( \prod_{j=1}^k \Lambda_j \right) \frac{t^k}{k!} |0\rangle_q \bigotimes_{j=1}^k \left( \sum_{\ell=0}^{L_j-1} |0\rangle_{c_j} \alpha_\ell \tilde{h}_\ell \right) \bigotimes_{j=k+1}^{\kappa} \underbrace{\left( \sum_{\ell=0}^{L_j-1} |0\rangle_{c_j} \alpha_\ell \right)}_{\Lambda_j |0\rangle_{c_j}} |\psi\rangle \\ &= |0\rangle \frac{1}{N_q \prod_{k=1}^{\kappa} \Lambda_k} \sum_{k=0}^{\kappa} \left( \prod_{j=1}^k \Lambda_j \right) \frac{t^k}{k!} \prod_{j=1}^k \left( \sum_{\ell=0}^{L_j-1} \alpha_\ell \tilde{h}_\ell \right) |\psi\rangle \\ &= \frac{1}{N_q} |0\rangle \sum_{k=0}^{\kappa} \frac{t^k}{k!} \prod_{j=1}^k \left( \sum_{\ell=0}^{L_j-1} \alpha_\ell \tilde{h}_\ell \right) |\psi\rangle = \frac{1}{N_q} |0\rangle U_{\mathbf{L}} |\psi\rangle = \Pi \mathcal{W}|0\rangle|\psi\rangle \end{aligned}$$

which is Eq. (10) and we see that  $N_q = s_{\mathbf{L}}$  as defined in Eq. (14).  $\square$

**Lemma 4.** The error of a single time step  $\delta_{\mathbf{L}}$  when using  $t_\infty = \log(2)/\Lambda$  can be bounded by

$$\delta_{\mathbf{L}}(t_\infty) \leq 2 - s_{\mathbf{L}}(t_\infty) =: \varepsilon_{\mathbf{L}}.$$

*Proof.* For easier notation, we first consider fully expanded orders and define

$$\begin{aligned} \tilde{\Delta}_n(t) &:= U_n(t) - U(t) \\ \varepsilon_n &:= s_\infty(t_\infty) - s_n(t_\infty) = 2 - s_n(t_\infty) \end{aligned}$$

and observe that

$$\begin{aligned} \|\tilde{\Delta}_n(t)\| &= \left\| \sum_{k=1}^n \frac{(-it)^k}{k!} \sum_{\ell_1, \dots, \ell_k=0}^{L-1} \alpha_{\ell_1} \dots \alpha_{\ell_k} h_{\ell_1} \dots h_{\ell_k} - \sum_{k=1}^{\infty} \frac{(-it)^k}{k!} \sum_{\ell_1, \dots, \ell_k}^L \alpha_{\ell_1} \dots \alpha_{\ell_k} h_{\ell_1} \dots h_{\ell_k} \right\| \\ &= \left\| \sum_{k=n+1}^{\infty} \frac{(-it)^k}{k!} \sum_{\ell_1, \dots, \ell_k=0}^{L-1} \alpha_{\ell_1} \dots \alpha_{\ell_k} h_{\ell_1} \dots h_{\ell_k} \right\| \\ &\leq \sum_{k=n+1}^{\infty} \frac{t^k}{k!} \sum_{\ell_1, \dots, \ell_k=0}^{L-1} \alpha_{\ell_1} \dots \alpha_{\ell_k} \underbrace{\|h_{\ell_1}\| \dots \|h_{\ell_k}\|}_1 \\ &= \left( 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{\ell_1, \dots, \ell_k=0}^{L-1} \alpha_{\ell_1} \dots \alpha_{\ell_k} \right) - \left( 1 + \sum_{k=1}^n \frac{t^k}{k!} \sum_{\ell_1, \dots, \ell_k=0}^{L-1} \alpha_{\ell_1} \dots \alpha_{\ell_k} \right) \\ &= s_\infty(t) - s_n(t) \end{aligned}$$

which means

$$\|\tilde{\Delta}_n(t_\infty)\| \leq s_\infty(t_\infty) - s_n(t_\infty) = \varepsilon_n.$$

Using these in our definition of  $\Delta_n$  yields

$$\begin{aligned} -\Delta_n(t_\infty) &= \tilde{\mathcal{A}}_n(t_\infty) - U(t_\infty) \\ &= \frac{3}{s_n(t_\infty)} U_n(t_\infty) - \frac{4}{s_n^3(t_\infty)} U_n(t_\infty) U_n^\dagger(t_\infty) U_n(t_\infty) - U(t_\infty) \\ &= \underbrace{\frac{3}{2 - \varepsilon_n}}_{=\frac{3}{2} + \frac{3\varepsilon_n}{4} + \mathcal{O}(\varepsilon_n^2)} (U + \tilde{\Delta}_n) - \underbrace{\frac{4}{(2 - \varepsilon_n)^3}}_{=\frac{1}{2} + \frac{3\varepsilon_n}{4} + \mathcal{O}(\varepsilon_n^2)} \overbrace{(U + \tilde{\Delta}_n)(U + \tilde{\Delta}_n)^\dagger(U + \tilde{\Delta}_n)}^{=U+2\tilde{\Delta}_n+U\tilde{\Delta}_n^\dagger U+\mathcal{O}(\tilde{\Delta}_n^2)} - U \\ &= \tilde{\Delta}_n \left( \frac{1}{2} - \frac{3\varepsilon_n}{4} \right) - U \tilde{\Delta}_n^\dagger U \left( \frac{1}{2} + \frac{3\varepsilon_n}{4} \right) + \mathcal{O}(\tilde{\Delta}_n^2) + \mathcal{O}(\varepsilon_n^2). \end{aligned}$$

Now we can finally bound the error to

$$\begin{aligned} \delta_n = \|\Delta_n(t_\infty)\| &\leq \left\| \tilde{\Delta}_n \left( \frac{1}{2} - \frac{3\varepsilon_n}{4} \right) \right\| + \left\| U \tilde{\Delta}_n^\dagger U \left( \frac{1}{2} + \frac{3\varepsilon_n}{4} \right) \right\| + \mathcal{O}(\varepsilon_n^2) \\ &\leq \frac{\varepsilon_n}{2} + \frac{\varepsilon_n}{2} + \mathcal{O}(\varepsilon_n^2) = \varepsilon_n + \mathcal{O}(\varepsilon_n^2), \end{aligned}$$

which straightforwardly extends to  $\delta_{\mathbf{L}}$  with  $\varepsilon_{\mathbf{L}}$  for partial orders.  $\square$

**Lemma 5.** *The error of  $r$  time steps  $\|U^r - \tilde{\mathcal{A}}_{\mathbf{L}}^r\|$  is bounded by  $r$  times the error of a single time step  $\delta_{\mathbf{L}} = \|\Delta_{\mathbf{L}}\| = \|U - \tilde{\mathcal{A}}_{\mathbf{L}}\|$ , up to order  $\delta_{\mathbf{L}}$ .*

*Proof.* We use the definition of  $\Delta = U - \tilde{\mathcal{A}}$  and substitute for  $\tilde{\mathcal{A}}$ .

$$\begin{aligned} \|U^r - \tilde{\mathcal{A}}_{\mathbf{L}}^r\| &= \|U^r - (U - \Delta)^r\| = \|U^r - U^r + \sum_{k=1}^r U^{k-1} \Delta U^{r-k} + \mathcal{O}(\Delta^2)\| \\ &\leq \sum_{k=1}^r \|U^{k-1} \Delta U^{r-k}\| + \|\mathcal{O}(\Delta^2)\| \leq \sum_{k=1}^r \|\Delta\| + \mathcal{O}(\delta^2) = r\delta + \mathcal{O}(\delta^2) \end{aligned}$$

$\square$

**Lemma 6.** *The logarithmic inverse error bound of a single time step for full orders  $\log(\varepsilon_n^{-1})$  scales like*

$$\log \frac{1}{\varepsilon_n} = \mathcal{O}(n \log n)$$

*Proof.* We can bound the residual of the Taylor series by

$$\begin{aligned} \varepsilon_n &= \sum_{k=n+1}^{\infty} \frac{\overbrace{(t_\infty \Lambda)^k}^{\log 2}}{k!} \\ &= \frac{\log^n 2}{n!} \sum_{k=1}^{\infty} \frac{n! \log^k 2}{(k+n)!} \\ &\leq \frac{\log^n 2}{n!} e^{\log 2} = \frac{2 \log^n 2}{n!} \end{aligned}$$

and use Stirling's approximation  $n! \leq e n^{n+1/2} e^{-n}$  to find

$$\varepsilon_n \leq \frac{2e^{-n} \log^n 2}{e n^{n+1/2}}$$

$$\log \varepsilon_n \leq \log 2 - n - n \log \log 2 - 1 - \left( n + \frac{1}{2} \right) \log n.$$

Therefore

$$\log \frac{1}{\varepsilon_n} = \mathcal{O}(n \log n).$$

□

**Corollary 7.** *Because the total complexity is of the order  $C_n = nL$ , the complexity when using full orders depending on the error bound  $\varepsilon$ , which we will call  $C_{\varepsilon, \text{full}}$  scales like*

$$C_{\varepsilon, \text{full}} = \mathcal{O} \left( \frac{L \log \frac{1}{\varepsilon}}{\log \log \frac{1}{\varepsilon}} \right)$$

*Proof.* We can use the inequality in Lemma 6

$$n < \log \frac{1}{\varepsilon_n}$$

to replace the  $n$  in the logarithm of Lemma 6 and find

$$n = \mathcal{O} \left( \frac{\log \frac{1}{\varepsilon_n}}{\log \log \frac{1}{\varepsilon_n}} \right)$$

and therefore

$$C_{\varepsilon, \text{full}} = \mathcal{O} \left( \frac{L \log \frac{1}{\varepsilon}}{\log \log \frac{1}{\varepsilon}} \right).$$

□

**Lemma 8.** *The bound for the logarithmic inverse error of the modified version  $\log \varepsilon_{\mathbf{L}}^{-1}$  scales like  $\mathcal{O} \left( \frac{C_{\mathbf{L}}}{L} \log \frac{C_{\mathbf{L}}}{L} \right) \leq \log \varepsilon_{\mathbf{L}}^{-1} < \mathcal{O}(C_{\mathbf{L}} \log C_{\mathbf{L}})$ , depending on the Hamiltonian.*

*Proof.* The left inequality follows immediately from the worst case that  $\alpha_0 = \alpha_1 = \dots = \alpha_L$ . In this case the modification is equivalent to the original method and Lemma 6 with  $n = C_{\mathbf{L}}/L$  holds.

In the other extreme case of  $\frac{\alpha_\ell}{\alpha_0} \rightarrow 0 \forall \ell \in \{1 \dots L\}$ , one term dominates the whole Hamiltonian, and adding  $h_0$  in some order of the expansion equates to adding that whole order, effectively reducing  $L$  to 1. Therefore Lemma 6 with  $L = 1$  defines an upper bound for the error scaling. □

**Corollary 9.** *The complexity of our modified version depending on the simulation error bound  $\varepsilon$ , which we call  $C_\varepsilon$ , is bounded by*

$$\mathcal{O} \left( \frac{\log \frac{1}{\varepsilon}}{\log \log \frac{1}{\varepsilon}} \right) < C_\varepsilon \leq \mathcal{O} \left( \frac{L \log \frac{1}{\varepsilon}}{\log \log \frac{1}{\varepsilon}} \right),$$

*Proof.* The same reasoning as in Corollary 7 applies to the bounds established in Lemma 8. □

**Corollary 10.** *The complexity of simulating for a time  $\tau$  with a total error bound  $\epsilon$  is bounded by*

$$\mathcal{O} \left( \frac{\Lambda \tau \log \frac{\Lambda \tau}{\epsilon}}{\log \log \frac{\Lambda \tau}{\epsilon}} \right) < C_\epsilon \leq \mathcal{O} \left( \frac{L \Lambda \tau \log \frac{\Lambda \tau}{\epsilon}}{\log \log \frac{\Lambda \tau}{\epsilon}} \right).$$

*Proof.* Simulating for a time  $\tau = r t_\infty = r \log(2)/\Lambda$  requires  $r$  steps. The error of a single step  $\varepsilon$  must therefore be  $\varepsilon = \epsilon/r$ . Substituting this in Corollary 9 and multiplying by the number of steps  $r = \mathcal{O}(\tau \Lambda)$  proves the claim. □