

# Quantum Science and Technology



## PAPER

# TE-PAI: exact time evolution by sampling random circuits

### OPEN ACCESS

RECEIVED  
22 June 2025

REVISED  
22 September 2025

ACCEPTED FOR PUBLICATION  
9 October 2025

PUBLISHED  
22 October 2025

Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Chusei Kiumi<sup>1,2,\*</sup> and Bálint Koczor<sup>1</sup>

<sup>1</sup> Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, United Kingdom

<sup>2</sup> Center for Quantum Information and Quantum Biology, Osaka University, 1-2 Machikaneyama, Toyonaka 560-0043, Japan

\* Author to whom any correspondence should be addressed.

E-mail: [c.kiumi.qiqb@osaka-u.ac.jp](mailto:c.kiumi.qiqb@osaka-u.ac.jp) and [balint.koczor@maths.ox.ac.uk](mailto:balint.koczor@maths.ox.ac.uk)

**Keywords:** quantum simulation, quantum algorithm, early fault-tolerant quantum computer

## Abstract

Simulating time evolution under quantum Hamiltonians is one of the most natural applications of quantum computers. We introduce TE-PAI, which simulates time evolution exactly by sampling random quantum circuits for the purpose of estimating observable expectation values at the cost of an increased circuit repetition. The approach builds on the Probabilistic Angle Interpolation (PAI) technique and we prove that it simulates time evolution without discretisation or algorithmic error while achieving shallow circuit depths with optimal scaling that saturates the Lieb–Robinson bound. Another significant advantage of TE-PAI is that it only requires executing random circuits that consist of Pauli rotation gates of only two kinds of rotation angles  $\pm\Delta$  and  $\pi$ , along with measurements. While TE-PAI is highly beneficial for NISQ devices, we additionally develop an optimised early fault-tolerant implementation using catalyst circuits and repeat-until-success teleportation, concluding that the approach requires orders of magnitude fewer T-states than conventional techniques, such as Trotterization—we estimate  $3 \times 10^5$  T states are sufficient for the fault-tolerant simulation of a 100-qubit Heisenberg spin Hamiltonian. Furthermore, TE-PAI allows for a highly configurable trade-off between circuit depth and measurement overhead by adjusting the rotation angle  $\Delta$  arbitrarily. We expect that the approach will be a major enabler in the late NISQ and early fault-tolerant periods as it can compensate circuit-depth and qubit-number limitations through an increased circuit repetition.

## 1. Introduction

Accurately modelling the time evolution of quantum systems is an important task but presents a significant challenge in classical computing. Thus, simulating quantum dynamics is regarded as one of the most promising applications of quantum computers [1, 2] and may provide an exponential speedup over classical computers. The simplest such approach, the Trotter–Suzuki decomposition [3, 4], approximates the time evolution through a relatively simple circuit that contains evolutions under the individual Hamiltonian terms. A drawback of the approach is that circuits may need to be quite deep to sufficiently suppress approximation errors. This discretisation error, also called the Trotter error, is inevitable with finite circuit depth and can be particularly daunting in, e.g. quantum chemistry applications [5–8]. These issues are further exacerbated when the aim is to simulate dynamics under time-dependent Hamiltonians, as we demonstrate below. Indeed, sophisticated quantum algorithms, such as linear combination of unitaries (LCU) [9–11], quantum signal processing [12, 13] or quantum walks [14, 15], can achieve a fundamentally improved circuit-depth scaling compared to Trotterisation; however, they require significant overheads in quantum resources.

We make significant progress and develop TE-PAI, which: (a) simulates ‘effectively exact’ time evolution on average; (b) requires executing only very simple circuits and performing measurements on them, i.e. it does not require advanced quantum resources such as ancillary qubits or controlled evolutions; (c) can naturally simulate time evolution under time-dependent Hamiltonians; (d) achieves shallow circuit depths with optimal scaling that saturates the Lieb–Robinson bound. TE-PAI is ‘effectively exact’ in the

sense that one needs to choose *a priori* a desired precision  $\epsilon$  to which the expectation value of the time-evolved observable is estimated, i.e. the level of statistical uncertainty (see appendix C for details). While algorithmic errors associated with time evolution are finite, they can be suppressed arbitrarily below  $\epsilon$  without increasing circuit depths, i.e. by committing  $O(\epsilon^{-1})$  classical pre-processing resources.

TE-PAI proceeds by constructing an unbiased estimator for the entire exact time-evolution superoperator—this allows us to estimate expectation values of time-evolved observables by sampling the output of quantum circuits and is thus compatible with advanced measurement techniques, such as classical shadows [16, 17] or Pauli grouping techniques [18, 19]. This immediately enables a wide range of applications, such as shadow spectroscopy for estimating energy gaps in the problem Hamiltonian [20]. TE-PAI can be combined straightforwardly with a broad range of further techniques that require the estimation of time-dependent correlators. Specifically, one can estimate the Loschmidt echoes  $\langle \psi(0) | \psi(t) \rangle$  and  $\langle \psi(t) | O | \psi(t') \rangle$  of fundamental importance, by estimating the Pauli  $Z$  expectation value on an ancilla qubit. This enables the estimation of ground and excited state energies via statistical phase estimation [21–23], the estimation of the density of states [24], as well as the estimation of expectation values of observables  $O$  in eigenstates of the problem Hamiltonian [25]. These techniques are expected to unlock some of the most important application areas of early fault-tolerant quantum computers (FTQCs), e.g. obtaining ground-state energies, correlation functions and molecular configurations, etc. However, we note that our approach is based on expected value estimation and is thus not compatible with conventional phase estimation or with other ‘single-shot’ techniques.

Previous work, such as qDRIFT [26–28], similarly use randomisation but the circuit depths are not independent of the approximation error. Prior works [29, 30] achieve some of the advantages of TE-PAI (exact evolution, comparable circuit depths), however, require more complex controlled circuits and are not compatible with advanced measurement techniques, such as classical shadows [16, 17] or Pauli grouping techniques [18, 19]. For instance, [30] requires estimating observables one-by-one using Hadamard tests and thus requires random circuits to be controlled on an ancillary qubit, as we detail in appendix A. In contrast, TE-PAI only requires executing simple random circuits and measuring their outputs, and thus benefits from almost unlimited compatibility with a broad range of applications and can be combined naturally with all error mitigation techniques, classical shadows, and randomised protocols that use time evolution as a subroutine, such as spectroscopy or estimating the density of states and beyond [16, 20, 24].

The main technical tool we exploit is Probabilistic Angle Interpolation (PAI) [31], which is particularly relevant for near-term and early fault-tolerant applications, where averaging over many circuit executions is required for expected-value measurements. While PAI increases the sampling cost, it probably achieves the least possible overhead. TE-PAI applies this mathematical formalism to time evolution circuits, generating random circuits with only two kinds of rotation angles,  $\pm\Delta$  and  $\pi$  as illustrated in figure 1; post-processing their outputs yields exact time-evolved expected values. Conventional trotterisation requires small gate angle settings that are proportional to the individual Hamiltonian terms. However, in real experiments the gate angles always have a finite precision as well as added coherent over or under rotations, which effectively leads to implementing the correct trotter evolution of a different Hamiltonian with slightly different Pauli coefficients—the effect of this may be radical near phase transitions. Our randomised approach effectively overcomes this limitation as the rotation angle  $\Delta$  can be chosen arbitrarily in our approach and can be updated according to our best knowledge of the true, experimentally calibrated rotation angle.

We prove that TE-PAI requires a number of gates proportional to the total simulation time  $T$  and the system size, thus saturating the Lieb–Robinson bound [32, 33]. A significant advantage of TE-PAI is that it offers a trade-off between sampling overhead and circuit depth, i.e. one can use NISQ-friendly shallow circuits at the cost of an increased sampling overhead. While these features are particularly important in the NISQ era, we construct optimised implementations for early FTQCs, whereby circuit-depth and width limitations will be similarly crucial.

Furthermore, we exploit that our approach works even if we use the same rotation angle in every gate in our time evolution circuits and develop an explicit, heavily optimised fault-tolerant implementation using catalyst circuits and repeat-until-success teleportation. We explicitly estimate fault-tolerant resource requirements for a benchmark problem and demonstrate that it may require significantly less magic states than prior techniques, e.g. we achieve orders of magnitude lower T-counts than conventional Trotterisation. As a result, our method imposes substantially lower fault-tolerance overheads, potentially enabling early practical demonstrations of quantum advantage in early FTQC.

This manuscript is structured as follows: We begin with a detailed description of the time-independent and time-dependent Trotter decompositions. In section 2, we derive the unbiased estimator of TE-PAI and prove that the sampling overhead and the expected number of gates in the circuit



It follows that achieving a precision  $\epsilon$  requires the following number of quantum gates in a Trotter circuit

$$\nu \leq \frac{1}{2}LT^2\|c\|_T^2\epsilon^{-1}. \quad (3)$$

Indeed, the additive Trotter error,  $\epsilon$ , can be reduced at the expense of proportionally increasing the circuit depth. In the next section we describe our algorithm which utilizes a probabilistic approach to generate random circuits from these Trotter-Suzuki circuit templates. In stark contrast, the circuit depth in our approach is independent of the precision, and  $N$  is a parameter that only affects the complexity of classical pre-processing as we detail in appendix C.

### 1.1.2. Time-dependent Hamiltonians

Building on the previous time-independent case, we extend the formalism to time-dependent Hamiltonians, which are crucial for accurately simulating complex quantum systems in practice that evolve under time-dependent interactions, such as in quantum control [38]. In particular, we consider a Hamiltonian  $H(t)$  whose decomposition coefficients  $c_k(t)$  are time-dependent as

$$H(t) = \sum_{k=1}^L c_k(t) h_k.$$

We will assume that  $c_k(t)$  are absolutely continuous functions of time. This assumption is necessary to bound the convergence rate in our proofs in appendices D and E. We then define their average  $\ell_1$ -norm as

$$\overline{\|c\|_1} := \frac{1}{T} \int_0^T \sum_{k=1}^L |c_k(t)| dt. \quad (4)$$

In this work, we consider the following discretised product approximation for the unitary evolution operator  $U(T)$  for a time-dependent Hamiltonian as

$$U(T) \approx e^{-iH(t_N)\frac{T}{N}} \dots e^{-iH(t_1)\frac{T}{N}}, \quad (5)$$

where  $H(t_j)$  represents the Hamiltonian at discrete time points via  $t_j = \frac{T}{N}j$  for  $j = 1 \dots N$  and  $\frac{T}{N}$ .

Finally, we define the Trotterised circuit for our time-dependent Hamiltonian as

$$U(T) \approx \prod_{j=1}^N \left( \prod_{k=1}^L e^{-ic_k(t_j)h_k\frac{T}{N}} \right).$$

The approximation assumes that the Hamiltonian remains constant within each small interval  $\frac{T}{N}$ , this approximation improves as  $N$  is increased, and error bounds have been reported in [37]. We also note that a generalisation of the above product formula can be found in [37] which has been derived by truncating the Magnus expansion [39] to the first order.

## 2. Unbiased estimators via TE-PAI

### 2.1. Unbiased estimator for general product formulas

In this section, we present details of our protocol that samples and post-processes measurement outcomes of random, shallow-depth quantum circuits in order to simulate effectively exact time evolution. By comparing equations (1) and (5) it is apparent that both time-dependent and time-independent product formulas are generally of the form (and similarly higher order product formulas can be written in this form) as

$$U = \prod_{j=1}^N \left( \prod_{k=1}^L R_k(\theta_{kj}) \right). \quad (6)$$

The gates above are Pauli rotation gates,  $R_k(\theta) := e^{-i h_k \theta / 2}$  and for time-dependent Hamiltonians their rotation angles are set to  $\theta_{kj} = 2c_k(t_j)\frac{T}{N}$ , which simplify to  $\theta_{kj} := \theta_k = 2c_k\frac{T}{N}$  for time-independent Hamiltonians. The parameter  $N$  controls the number of Trotter steps, and by increasing  $N$ , we can approximate the desired time evolution operator with arbitrary accuracy via statement (1).

We use the *Probabilistic Angle Interpolation* (PAI) technique [31], which we summarise in appendix B, and which builds on the observation that estimating an expected value of an observable requires a quantum circuit to be run and measured many times. At each circuit run, PAI randomly replaces the angle  $\theta_{kj}$  in the rotation gate  $R_k(\theta_{kj})$  with one of only three discrete rotation angles  $0$ ,  $\text{sign}(\theta_{kj})\Delta$ , or  $\pi$ . We will denote these three gate variants as

$$A = I, \quad B_{kj} = R_k(\text{sign}(\theta_{kj})\Delta), \quad C_k = R_k(\pi). \quad (7)$$

We note that one can choose a uniquely different  $\Delta_{kj}$  specifically for each rotation angle  $\theta_{kj}$ , however, for ease of notation we assume that one global  $\Delta$  is chosen such that  $|\theta_{kj}| \leq \Delta < \pi$  for all  $k \in 1, \dots, L$ —this choice will significantly reduce resources required for fault-tolerant implementations as we will detail below.

The crucial observation that TE-PAI exploits is that in the limit  $N \rightarrow \infty$ , where the Trotter circuit approaches exact evolution, the rotation angles  $\theta_{kj}$  become vanishingly small. Thus, with probability nearly equal to 1, we almost always choose the first gate variant, the identity operation (see appendix D). As we prove below, this ultimately guarantees that the total circuit depth remains finite even in the limit  $N \rightarrow \infty$  and thus the parameter  $N$  only influences the complexity of classical pre-computation.

We review in detail in appendix B that the PAI approach builds on the fact that the superoperator representation  $\mathcal{R}_k(\theta_{kj})$  of each unitary gate  $R_k(\theta_{kj})$  can be decomposed analytically as

$$\mathcal{R}_k(\theta_{kj}) = \gamma_1(|\theta_{kj}|)\mathcal{A} + \gamma_2(|\theta_{kj}|)\mathcal{B}_{kj} + \gamma_3(|\theta_{kj}|)\mathcal{C}_k, \quad (8)$$

where  $\mathcal{A}$ ,  $\mathcal{B}_{kj}$  and  $\mathcal{C}_k$  are superoperator representations of the unitary gates in equation (7) which generally act isomorphically via conjugation, e.g.  $\mathcal{R}_k \text{vec}[\rho] = \text{vec}[R_k \rho R_k^\dagger]$ . The coefficients  $\gamma_l(\theta)$  are provided explicitly as trigonometric functions in appendix B.

Focusing on a single gate element  $\mathcal{R}_k(\theta_{kj})$ , our classical pre-processing algorithm randomly selects one of the three discrete gate variants as  $\hat{\mathcal{D}}_l \in \{\mathcal{A}, \mathcal{B}_{kj}, \mathcal{C}_k\}$  according to the probabilities  $p_l = |\gamma_l(\theta_{kj})| / \|\gamma(\theta_{kj})\|_1$  for  $l \in \{1, 2, 3\}$  in order to sample the unbiased estimator of the desired, continuous-angle gate  $\mathcal{R}_k(\theta_{kj})$  as

$$\hat{\mathcal{R}}_k(\theta_{kj}) = \|\gamma(|\theta_{kj}|)\|_1 \text{sign}[\gamma_l(|\theta_{kj}|)] \hat{\mathcal{D}}_l. \quad (9)$$

Koczor *et al* [31] analytically proved that this choice of the three discrete angle settings minimises the measurement overhead characterised by  $\|\gamma(|\theta_{kj}|)\|_1$ , confirming numerical optimisation results of [40].

Replacing each gate in the product formula in equation (6) with the above unbiased estimator then allows us to construct an unbiased estimator for the entire time evolution operator. As we now summarise, this is a direct consequence of statement 2 of [31].

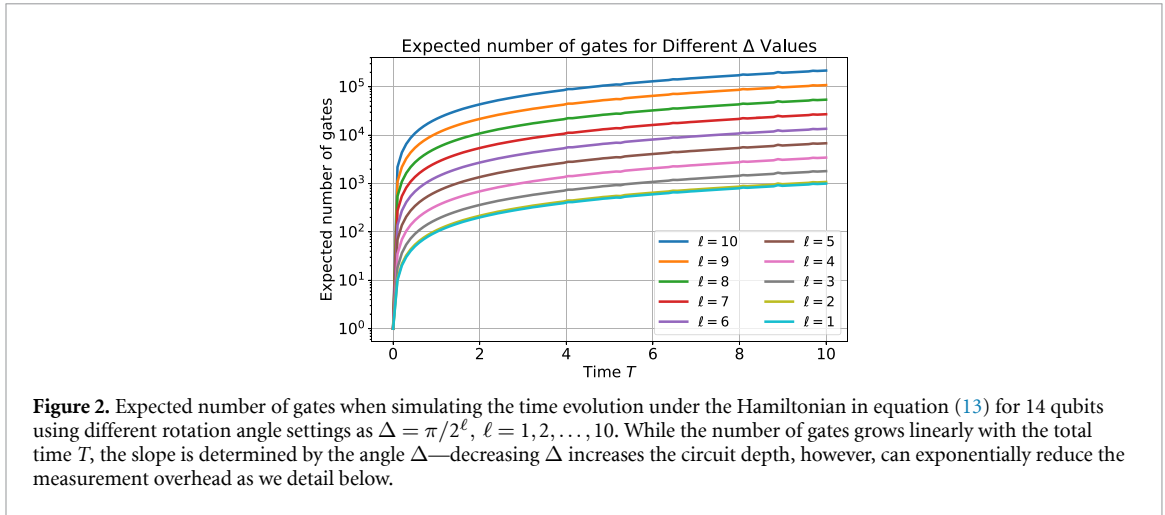
**Statement 2.** We obtain an unbiased estimator of the superoperator representation of the entire product formula in equation (6) as

$$\hat{\mathcal{U}} = \prod_{j=1}^N \left( \prod_{k=1}^L \hat{\mathcal{R}}_k(\theta_{kj}) \right), \quad (10)$$

using the unbiased estimators  $\hat{\mathcal{R}}_k(\theta_{kj})$  of the individual continuous-angle rotations from equation (9). Specifically, the mean value of the estimator is  $\mathbb{E}[\hat{\mathcal{U}}] = \mathcal{U}$ . The classical computational complexity of generating  $N_s$  random circuits is  $O(NLN_s)$ .

Please refer to appendix B for a detailed derivation. Figure 1 shows an example of a random circuit generated using the PAI approach that simulates the time evolution under the Hamiltonian equation (13) for 5-qubit. Finally our protocol is summarised as follows.

- Take a quantum circuit  $\mathcal{U}$  of the form of equation (6) which implements a product formula for simulating the time evolution under the input Hamiltonian  $\mathcal{H}$  with parameters  $N$  and  $T$ .
- Generate  $N_s$  random circuits by randomly replacing gates in the circuit  $\mathcal{U}$  with fixed rotation angles of  $\pm\Delta$  and  $\pi$  according to the PAI protocol in statement 2.
- Execute all random circuits and in post-processing multiply all measurement outcomes with their corresponding prefactor  $\prod_{j=1}^N \prod_{k=1}^L \|\gamma(|\theta_{kj}|)\|_1 \text{sign}[\gamma_l(|\theta_{kj}|)]$  where the index  $l := l_{kj}$  is chosen randomly for each gate in the circuit as  $l \in \{1, 2, 3\}$ . When classical shadows are estimated then the expected value from each snapshot needs to be multiplied by this factor as detailed in [16].



## 2.2. Gate count in the random circuits

For ease of notation, in the following we specifically consider first-order Trotter circuits in equations (1) and (5), but our proofs apply to any higher-order product formula. In the standard first-order Trotter approach, the number of gates  $\nu = NL$  is directly proportional to  $N$ . In contrast, TE-PAI generates circuits randomly, and the number of gates is thus formally a random variable. We now prove that, as we increase  $N$ , the mean value  $\mathbb{E}[\nu]$  asymptotically approaches a constant.

**Theorem 1.** *The expected number of gates  $\mathbb{E}(\nu)$  can be approximated in terms of*

$$\nu_\infty := \lim_{N \rightarrow \infty} \mathbb{E}(\nu) = \csc(\Delta) (3 - \cos \Delta) \overline{\|c\|_1} T$$

up to an error term as  $\mathbb{E}(\nu) = \nu_\infty + O(N^{-1})$ . Furthermore, the variance of the gate count satisfies the same scaling:  $\text{Var}[\nu] = \nu_\infty + O(N^{-1})$ . The asymptotic gate count is lower bounded as

$$\nu_\infty \geq \overline{\|c\|_1} T 2\sqrt{2}. \quad (11)$$

This bound is saturated when using the large angle  $\Delta = 2 \arctan(1/\sqrt{2}) \approx 0.392\pi$ . In the special case of time-independent Hamiltonians the same result holds up to formally replacing  $\overline{\|c\|_1} \equiv \|c\|_1$ .

We explicitly bound the leading constant factors in the error term in  $\mathbb{E}(\nu) = \nu_\infty + O(N^{-1})$  in our proof in appendix D. We thus find that for a constant rotation angle  $\Delta$ , TE-PAI saturates the Lieb–Robinson bound [32, 33], which determines a fundamental bound on the speed at which local information can spread due to time evolution under local interactions. However, we will later prove that in practice,  $\Delta$  needs to scale with both  $T$  and the system size to avoid an exponential increase in the sample complexity.

In addition to the above mean value, we also characterise the distribution of the number of gates in the circuit.

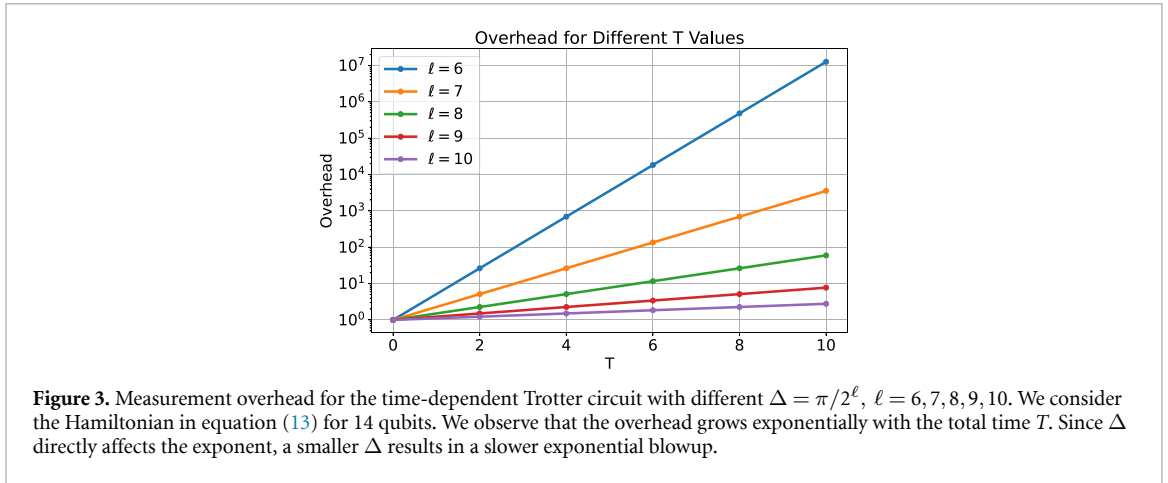
**Lemma 1.** *The distribution of the number of gates in the circuit approaches a normal distribution  $\mathcal{N}(\nu_\infty, \nu_\infty)$  as  $N \rightarrow \infty$ , where  $\nu_\infty$  is the mean value from theorem 1.*

As illustrated in figure 2, the expected number of gates for the time-independent Trotter circuit grows linearly with the total time  $T$ .

Additionally, the figure demonstrates that decreasing  $\Delta$  results in an increased number of gates, however, as we will see below, it also decreases the measurement overhead exponentially.

## 2.3. Measurement overhead

In TE-PAI, we randomly replace the continuous-angle gates with three discrete gate variants: the third gate variant has a very low associated probability for small  $\Delta$ , however, when it does get selected then any measured observable is multiplied with a negative sign. This negative sign leads to an increase in the variance of the expectation value of the observable being estimated. Thus, in order to estimate the observable expected value to the same precision as with an infinitely deep Trotter circuit, one needs to



perform an increased number of measurements. We detail in appendix B that this measurement overhead is upper bounded by the following factor as

$$\|g\|_1 := \prod_{j=1}^N \prod_{k=1}^L \|\gamma(|\theta_{kj}\rangle)\|_1. \quad (12)$$

Now we prove that the measurement overhead  $\|g\|_1$  converges to a constant as we increase  $N$ .

**Theorem 2.** *We bound the number of shots  $N_s$  required to achieve a specified precision  $\epsilon$  in estimating time-evolved expectation values. The number of circuit repetitions in TE-PAI is upper bounded as*

$$N_s \leq \|g\|_1^2 / \epsilon^2,$$

whereas having access to an infinitely deep Trotter circuit results in the upper bound  $N_s \leq \epsilon^{-2}$ . The overhead determined by  $\|g\|_1^2$  can be approximated via

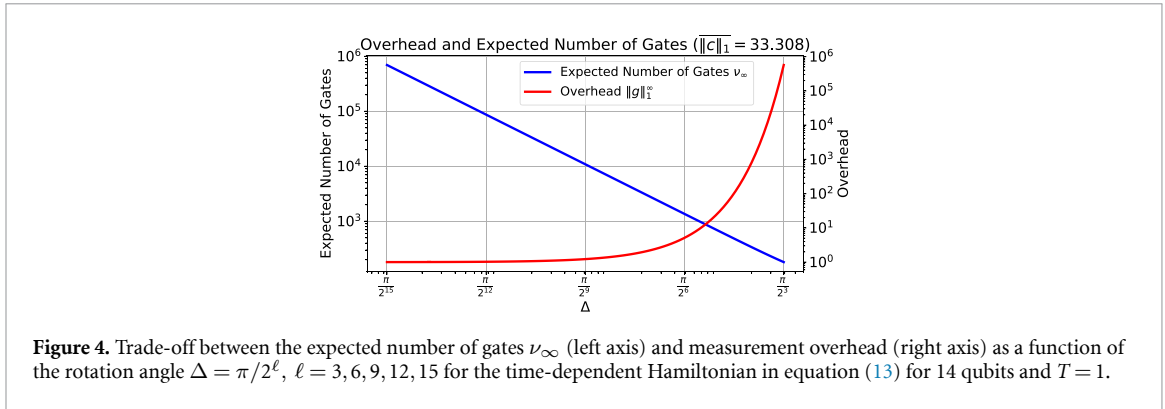
$$\|g\|_1^\infty := \exp \left[ 2 \overline{\|c\|_1} T \tan \left( \frac{\Delta}{2} \right) \right]$$

up to an term as  $\|g\|_1 = \|g\|_1^\infty + O(N^{-1})$ . In the special case of time-independent Hamiltonians the measurement overhead simplifies via  $\overline{\|c\|_1} \equiv \|c\|_1$ .

We explicitly derive and bound the leading constant factors in the above error term in our proof in appendix E. We note that a variant of PAI was developed in [40] that introduces a trade-off parameter  $\lambda$  that allows to continuously interpolate between the unbiased PAI estimator (exact estimator with measurement overhead  $\|g\|_1$ ) and an approximate, biased estimator which has no measurement overhead at all. The numerical approach of [40] can be used immediately for reducing the above measurement overhead at the cost of introducing a bias, however, for ease of notation in the present work we focus on the exact, unbiased implementation.

We illustrate in figure 3, that the measurement overhead of our unbiased estimator grows exponentially with the total time  $T$  but decreasing  $\Delta$  results in a slower exponential blowup. Furthermore, in figure 4, we illustrate the trade-off between the measurement overhead and the expected number of gates at different rotation angles  $\Delta$ . As we increase  $\Delta$ , we decrease the circuit depth but also increase the measurement overhead. While the expected number of gates  $\nu_\infty$  is a constant that is independent of  $N$ , decreasing the rotation angle  $\Delta$  increases  $\nu_\infty$  and ultimately can yield to divergence in the limit  $\lim_{\Delta \rightarrow 0} \mathbb{E}(\nu) = \infty$ . In a quantum computer where the quantum gates are not perfect—for example, noisy physical operations or logical operations in an early fault-tolerant machine—the circuit depth needs to be chosen such that  $p_{\text{err}}^{-1} \propto \nu$  to ensure that error-mitigation techniques can be used effectively. This desired circuit depth can be achieved by setting the corresponding  $\Delta$  angle.

Both the measurement overhead and the circuit depth of TE-PAI depends implicitly on the number of qubits, through the dependence of the norm  $\overline{\|c\|_1}$  on the number of qubits  $n$ . For example, for the Heisenberg spin model considered later in this manuscript,  $\overline{\|c\|_1} \in O(n)$ , while in quantum chemistry a pessimistic worst-case bound is  $\overline{\|c\|_1} \in O(n^4)$ . However, advanced techniques such as tensor hypercontraction can significantly reduce this scaling [41]. The complexity of other quantum simulation methods,



**Figure 4.** Trade-off between the expected number of gates  $\nu_\infty$  (left axis) and measurement overhead (right axis) as a function of the rotation angle  $\Delta = \pi/2^\ell$ ,  $\ell = 3, 6, 9, 12, 15$  for the time-dependent Hamiltonian in equation (13) for 14 qubits and  $T = 1$ .

such as qubitization and qDRIFT, similarly depends implicitly on the number of qubits via the norm  $\|c\|_1$ .

Finally, we consider fixing the measurement overhead as a constant and establish how the rotation angle  $\Delta$  scales with the system size and total simulation time.

**Remark 1.** We introduce a trade-off parameter  $Q$  that governs the trade-off between circuit depth and the measurement overhead. By using a rotation angle  $\Delta = 2 \arctan\left(\frac{Q}{2\|c\|_1 T}\right)$ , we achieve a constant overhead of  $\exp(Q)$  and obtain the number of gates as

$$\nu_\infty = \frac{2\left(\|c\|_1 T\right)^2}{Q} + Q \leq \frac{4\left(\|c\|_1 T\right)^2}{Q}.$$

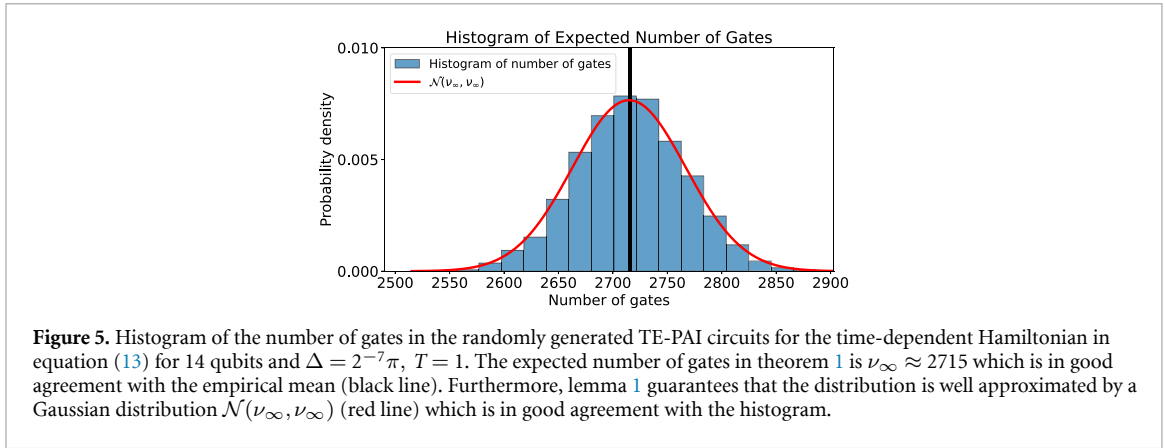
The upper bound above is due to the fact that  $Q \leq \|c\|_1 T \sqrt{2}$ , via the lower bound on  $\nu_\infty$  in equation (11).

In practice one would choose  $Q \geq 1$  given the measurement overhead  $\exp(1)$  is still very reasonable. Let us now compare the number of gates and the total time complexity of TE-PAI to similar techniques.

**Number of gates:** The parameter  $Q$  allows us to have a constant measurement overhead at the cost of the number of gates increasing quadratically with the system size and with the time depth  $T$ . This scaling is the same as in the case of first-order Trotterisation whereby  $\nu \in O(T^2/\epsilon)$  in equation (C1), however, the crucial difference is that the constant factor  $\epsilon^{-1}$  in Trotterisation is replaced here with a controllable hyperparameter  $Q^{-1} \leq 1$  which in practice is many orders of magnitude smaller. Thus, we expect TE-PAI to require orders of magnitude fewer gates than first-order Trotterisation.

**Time complexity:** We now estimate the end-to-end time complexity of TE-PAI. The number of gates in a single circuit is  $O(T^2 Q^{-1})$ , and estimating an observable to precision  $\epsilon$  requires  $O(\epsilon^{-2} \exp Q)$  repetitions (we assume the circuits are repeated serially and therefore TE-PAI does not require a space overhead). The end-to-end time complexity of TE-PAI (or equivalently its space-time volume) therefore scales as  $O(T^2 \epsilon^{-2} Q^{-1} \exp[Q])$ . We can compare this to a first-order Trotter circuit from which the expected value is extracted using amplitude estimation in which case the Trotter circuit is repeated coherently  $O(\epsilon^{-1})$  times leading to a total time complexity  $O(T^2 \epsilon^{-2})$  (the space-time volume of this approach scales similarly given amplitude estimation requires little space overhead).

In conclusion, the time complexity of extracting a time-evolved expected value in TE-PAI is comparable to using first-order Trotterisation in combination with amplitude estimation. However, TE-PAI has a number of significant advantages. First, TE-PAI requires only shallow circuits, making it feasible to run when coherence time or code distance is limited, whereas amplitude amplification requires significantly deeper circuits. Second, while amplitude estimation estimates observables one-by-one, we detail below that TE-PAI is compatible with advanced measurement techniques, including classical shadows and thus allows simultaneous estimation of many observables. Third, certain platforms, such as silicon qubit technologies, enable the fabrication of many independent QPUs on a single chip with no or limited communication between QPUs. TE-PAI can be fully parallelized, with the sampling task distributed across many QPUs, leading to a proportional reduction of the runtime. In contrast, some amplitude estimation variants can be parallelised, but only to a more limited extent [42]. Finally, TE-PAI only requires the implementation of a single type of non-Clifford rotation, specifically a single-qubit rotation with angle  $\Delta$ , which allows us to design a particularly efficient fault-tolerant implementation below.



Note that, in terms of gate complexity, truncated Taylor-series LCU methods scale as  $\mathcal{O}\left(\|c\|_1 T \frac{\log(1/\epsilon)}{\log \log(1/\epsilon)}\right)$  [9–11], while qubitization, QSP, and QSVT scale as  $\mathcal{O}(\|c\|_1 T + \log(1/\epsilon))$  [12, 13, 43], but these techniques are fundamentally different and require deep coherent circuits that are not feasible to execute on current or early-FTQC devices unlike TE-PAI.

### 3. Numerical demonstrations

#### 3.1. Heisenberg spin model benchmark

We consider the benchmarking task of simulating the spin-ring Hamiltonian as

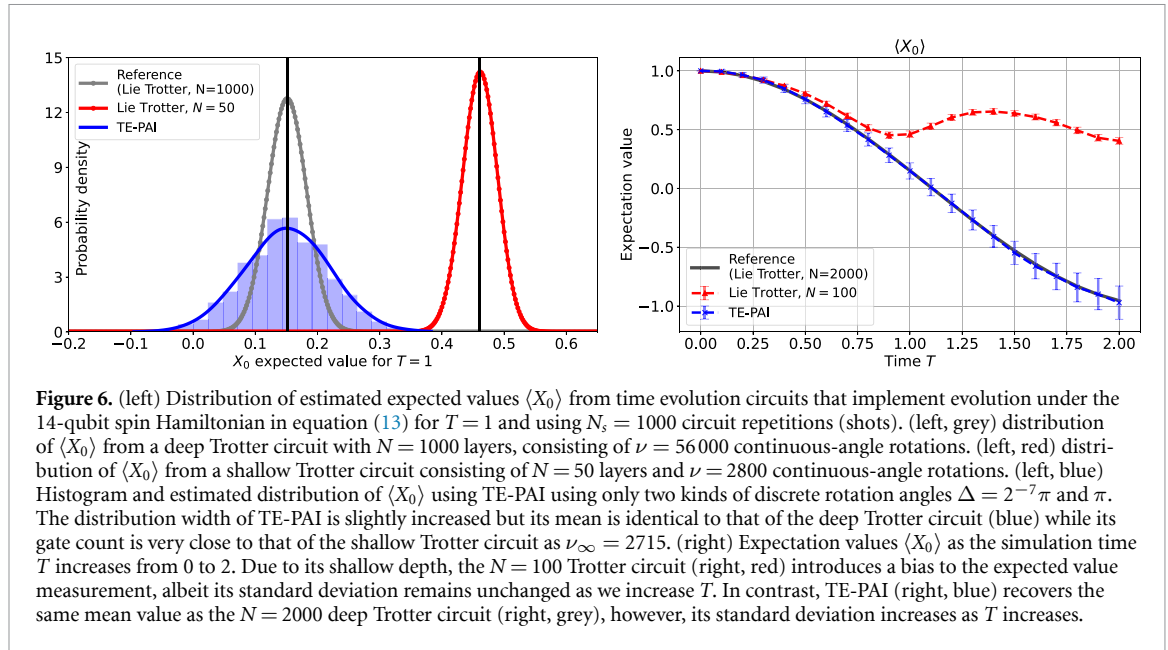
$$\mathcal{H} = \sum_{k \in \text{ring}(N)} \omega_k Z_k + J(t) \vec{\sigma}_k \cdot \vec{\sigma}_{k+1}, \quad (13)$$

where we choose time-dependent coupling terms  $J(t) = \cos(99\pi t)$  and the parameters  $\omega_k$  are chosen uniformly randomly within the range  $[-1, 1]$ . This model is representative of problems considered in condensed matter physics for studying many-body localisation. These problems could be effectively explored using early quantum computers and may be hard to simulate classically for large numbers of qubits [44, 45]. For the present demonstration, we use  $n = 14$  qubits and due to the periodicity of the Hamiltonian,  $\|c\|_1 \approx 33.30$  is constant for integer evolution times  $T = 1, 2, \dots$ , and we choose  $T = 1$  and  $\Delta = 2^{-7}\pi$ . We generate random circuits using TE-PAI for a product formula of  $N = 1000$  layers and a total evolution time of 1—we chose a relatively low value of  $N$  as in our demonstrations we will use a relatively low number of shots  $N_s = 1000$  and therefore shot noise will dominate over residual algorithmic errors (while indeed  $N$  can be increased without requiring more quantum resources).

First, we present a histogram in figure 5 that estimates the distribution of the number of gates from  $N_s = 1000$  different randomly generated TE-PAI circuits. The expected number of gates via theorem 1 is  $\nu_\infty \approx 2715$  which shows good agreement with the empirical mean in our histogram (black line). Furthermore, as predicted by lemma 1, the histogram agrees well with a Gaussian distribution  $\mathcal{N}(\nu_\infty, \nu_\infty)$ .

Second, we execute the TE-PAI circuits to estimate expected values  $\langle X_0 \rangle$  and compare them to conventional Lie–Trotter circuits. In figure 6(left) we report the distribution of expected values estimated using  $N_s = 1000$  shots: while TE-PAI in figure 6(left, blue) has a slightly increased distribution width, its mean value matches exactly the mean value of an arbitrarily deep Trotter circuit in figure 6(left, grey). Furthermore, with small measurement overhead of  $\|g\|_1^\infty \approx 2.15$ , TE-PAI requires only a small number of gates comparable to a shallow Trotter circuit with  $N = 50$ —which shallow circuit introduces a significant bias due to significant algorithmic errors, as shown in figure 6(left, red).

Finally, we generate a family of TE-PAI circuits for an increasing total simulation time  $T$  and compare the expected values  $\langle X_0 \rangle(T)$  to shallow and deep Trotter circuits in figure 6(right). Since this plot considers  $T = 0 \dots 2$ , we doubled the Trotter step numbers to  $N = 2000$  for the deep Trotter circuit and  $N = 100$  for the shallow Trotter circuit. Expected values estimated from shallow Trotter circuits have relatively low statistical uncertainty figure 6(right, red error bars) throughout the evolution, however, suffer from a significant bias, i.e. red curve is significantly off from the reference deep time evolution circuit figure 6(right, grey). In contrast, TE-PAI (blue line) achieves the same mean value as the deep Trotter circuit, however, its statistical uncertainty increases with the simulation time.



**Figure 6.** (left) Distribution of estimated expected values  $\langle X_0 \rangle$  from time evolution circuits that implement evolution under the 14-qubit spin Hamiltonian in equation (13) for  $T = 1$  and using  $N_s = 1000$  circuit repetitions (shots). (left, grey) distribution of  $\langle X_0 \rangle$  from a deep Trotter circuit with  $N = 1000$  layers, consisting of  $\nu = 56000$  continuous-angle rotations. (left, red) distribution of  $\langle X_0 \rangle$  from a shallow Trotter circuit consisting of  $N = 50$  layers and  $\nu = 2800$  continuous-angle rotations. (left, blue) Histogram and estimated distribution of  $\langle X_0 \rangle$  using TE-PAI using only two kinds of discrete rotation angles  $\Delta = 2^{-7}\pi$  and  $\pi$ . The distribution width of TE-PAI is slightly increased but its mean is identical to that of the deep Trotter circuit (blue) while its gate count is very close to that of the shallow Trotter circuit as  $\nu_\infty = 2715$ . (right) Expectation values  $\langle X_0 \rangle$  as the simulation time  $T$  increases from 0 to 2. Due to its shallow depth, the  $N = 100$  Trotter circuit (right, red) introduces a bias to the expected value measurement, albeit its standard deviation remains unchanged as we increase  $T$ . In contrast, TE-PAI (right, blue) recovers the same mean value as the  $N = 2000$  deep Trotter circuit (right, grey), however, its standard deviation increases as  $T$  increases.

### 3.2. Quantum chemistry benchmark

The complexity of TE-PAI depends on the coefficient  $\ell_1$ -norm  $\overline{\|c\|_1}$  rather than on the number of Hamiltonian terms  $L$ . This suggests that TE-PAI could be advantageous for quantum chemistry problems with large  $L$ , where standard Trotter methods require prohibitively large gate counts per Trotter step. An important example is the simulation of the FeMoCo active-space Hamiltonian. TE-PAI is therefore expected to play a role in enabling quantum advantage in challenging quantum chemistry applications.

As a simple demonstration, we consider a time-independent chemistry simulation of a linear chain of  $N_H = 6$  hydrogen atoms in the STO-6G minimal basis with an interatomic spacing of 2.0 Bohr, described by the corresponding electronic Hamiltonian:

$$H = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \quad (14)$$

where  $a_p^\dagger$  and  $a_q$  are fermionic creation and annihilation operators, and  $h_{pq}$ ,  $h_{pqrs}$  denote one- and two-electron integrals obtained from a restricted Hartree–Fock calculation. Mapping to qubits is performed using the Jordan–Wigner transformation, which yields a Pauli expansion acting on twelve qubits (corresponding to twelve spin orbitals).

Since orbital occupation numbers,

$$n_p = a_p^\dagger a_p = \frac{1 - Z_p}{2},$$

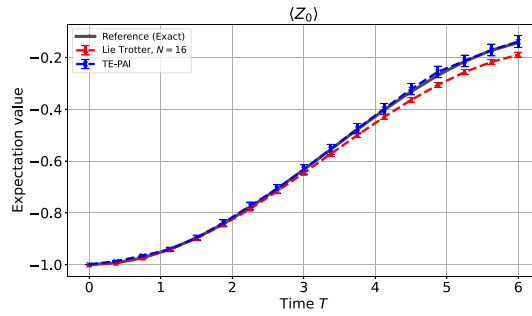
are an important observable in quantum chemistry, we compute the single-qubit expectation value  $\langle Z_0(t) \rangle$ , which is directly related to the occupation of the first spin orbital.

Figure 7 shows the results of simulating time evolution under the  $H_6$  Hamiltonian with 919 Pauli terms, starting from the initial state  $|101001010101\rangle$  and using  $N_s = 10000$  circuit repetitions (shots), as the simulation time  $T$  increases from 0 to 6. We compare the estimated expectation values  $\langle Z_0(t) \rangle$  with an exact simulation obtained by directly exponentiating the Hamiltonian, a shallow Trotter circuit with  $N = 16$  layers requiring  $\nu = 14704$  continuous-angle rotations that introduces systematic bias, and TE-PAI with  $\nu_\infty = 14138$  gates and fixed rotation angle  $\Delta = 2^{-8}\pi$ , implemented using random circuits corresponding to a product formula with  $N = 1000$  layers. In this setting, TE-PAI achieves the same unbiased result as the exact simulation with a gate cost comparable to that of the shallow Trotter circuit.

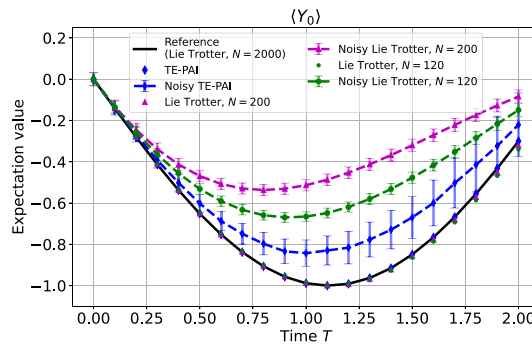
In chemistry simulations typically a relatively high precision, the chemical accuracy  $\propto 10^{-3}$  is required which is challenging to meet using shallow Trotter circuits. TE-PAI, on the other hand, reproduces the exact result, with the only remaining error coming from sampling noise, which can be reduced by increasing the number of shots.

### 3.3. Noisy implementation

Before fault tolerance is achieved, one needs to resort to noisy physical gates to execute circuits in the NISQ era. This poses limitations on the achievable circuit depths, as the total number of noisy gates



**Figure 7.** Expectation values  $\langle Z_0 \rangle$  from time evolution under the  $H_6$  Hamiltonian with 919 Pauli terms, starting from the initial state  $|1010010101\rangle$  and using  $N_s = 10\,000$  circuit repetitions (shots), as the simulation time  $T$  increases from 0 to 6. (grey) Exact simulation obtained by directly exponentiating the Hamiltonian. (red) Shallow Trotter circuit with  $N = 16$  layers, consisting of  $\nu = 14\,704$  continuous-angle rotations, which introduces a systematic bias in the estimated values. (blue) TE-PAI simulation using discrete rotation angles with  $\Delta = \pi/2^8$ , which recovers the same mean as the exact reference while requiring only  $\nu_\infty = 14\,138$  gates; however, its standard deviation increases with  $T$ .



**Figure 8.** Expected value  $\langle Y_0 \rangle$  after time evolution using 1000 circuit repetitions (shots) in a 7-qubit spin-ring model from section 3 using noisy quantum gates. Our reference is a noise-free Lie–Trotter simulation with  $N = 2000$  Trotter steps (grey) consisting of  $\nu = 56\,000$  parameterized gates; Using a shallower Trotter circuit as  $N = 120$  layers consisting of  $\nu = 3360$  noisy parameterized gates (green) achieves a smaller bias than using  $N = 200$  layers (magenta) consisting of  $\nu = 5600$  noisy parameterized gates. However, TE-PAI (blue) achieves the smallest bias as it uses a fewer number of noisy quantum gates at the expense of an increased statistical uncertainty (increasing error bars).

is typically restricted to a small constant multiple of the inverse average gate error rate. Compared to Trotterisation, TE-PAI has the significant advantage that the circuit depth can be reduced without introducing bias. Here we demonstrate the improved robustness of TE-PAI against gate noise and consider a noise model where each gate is followed by depolarisation that acts on the same qubit(s) as the gate itself. We assume a typical two-qubit gate error probability of  $p_2 = 10^{-3}$  and for single-qubit gates we assume an order of magnitude lower error probability of  $p_1 = 10^{-4}$ .

In figure 8, we repeat simulations of the spin-ring model defined in section 3 but assuming a noisy 7-qubit system and a rotation angle  $\Delta = \frac{\pi}{2^6}$  which yields an expected number of gates of 1364 at  $T = 2$ . We consider Trotterisation using  $N = 100$  (green) and  $N = 200$  layers (magenta). For both  $N = 100$  and  $N = 200$ , the Trotter step sizes are larger than those of the shallow Trotter circuit in figure 6, but they are still small enough that the Trotter error cannot be ignored in our noiseless simulations. Despite the small number of steps, the resulting circuits remain highly susceptible to noise. We also observe that increasing the number of Trotter layers introduces a more significant bias due to the increased number of noisy gates. In contrast, TE-PAI achieves a smaller bias, i.e. blue is closer to the reference simulation (grey), as it requires fewer noisy quantum gates. This configuration of  $\Delta$  enables us to flexibly reduce the circuit depth, thereby enhancing its resilience to noise compared with Trotter circuits.

Additionally, to further reduce the effect of gate noise, TE-PAI can naturally be combined with quantum error mitigation [46] and can also be combined with classical shadows as detailed in [16]. Furthermore, TE-PAI estimates an expectation value by executing a large number of structurally radically different random circuits; such randomisation protocols have been shown to scramble local gate noise to global depolarising noise—with theoretical proofs for global random circuits [47] and numerical evidence for shallow structured circuits [48]—which allows for very simple and effective error mitigation by global rescaling.

**Table 1.** Resource estimates for the fault-tolerant simulation of a 100-qubit spin-ring Hamiltonian in equation (13), with  $L = 400$  Hamiltonian terms,  $T = 1$ ,  $\|c\|_1 = 241.3$ , and  $\Delta = \frac{\pi}{28}$ . The expected number of  $R_G(\pm\Delta)$  gates is approximately 39 328, with a measurement overhead of 19.32. For the Trotter circuit, we use the direct synthesis method assuming  $N = 10\,000$  trotter steps. While we assumed limited storage space, the efficiency of the Hamming phasing approach would asymptotically for large number of storage qubits approach that of the catalyst towers.

	Trotter circuit	TE-PAI, direct synthesis	TE-PAI, Hamming weight phasing	TE-PAI, towers
T-gates	356 000 000	2438 336	1880 980	298 647
storage qubits	—	—	63	63
ancilla qubits	—	—	56	60

## 4. Fault-tolerant resource estimation

TE-PAI implements time evolution exactly by averaging outputs of random circuits. A significant advantage compared to, e.g. Trotterisation, is that our circuits only have two kinds of rotation angles as  $\pm\Delta$  and  $\pi$ , and, as we demonstrate in the following, this significantly reduces overall non-Clifford resources required for fault-tolerant implementations.

In particular, early generations of FTQCs will likely be limited by the number of logical qubits and by the achievable circuit depths. Here, we perform resource estimation for a typical example of simulating the time evolution under a 100-qubit spin Hamiltonian (which we introduced in section 3). This Hamiltonian has  $L = 400$  terms, and we fix  $T = 1$ ,  $\|c\|_1 = 241.3$ . The expected number of rotation gates using  $\Delta = \frac{\pi}{28}$  is approximately 39 328 and the measurement overhead is then about 19.32. We set the Trotter step to  $N = 10^4$  because, by equation (C1) in appendix C, this choice guarantees a Trotter error to be  $\epsilon_T \leq 0.1$ . However, such worst-case bounds are typically loose by about an order of magnitude [36], so we expect a error on the order of  $10^{-2}$  in practice. Table 1 summarises the resource estimation, indicating the total number of T-gates required for a single circuit in the Trotter and TE-PAI methods using different approaches.

In appendix G, we detail how our random circuits can be compiled into a sequence of Clifford gates and discrete angle, single-qubit  $Z$  rotation gates  $R_Z(\Delta)$ . Thus the only non-Clifford resource we require are single-qubit  $Z$  rotation gates, all with the same rotation angle. In fault-tolerant quantum computing (FTQC), Clifford gates are relatively cheaper and less error prone than non-Clifford resources. Thus, our focus in this section is to minimize the implementation cost of our method by efficiently implementing the  $R_Z(\Delta)$  rotations.

Fault-tolerant machines are emerging; however, there is a range of different hardware platforms with radically different qubit technologies. For example, solid-state platforms (superconducting qubits and quantum dots) have fast gates but relatively noisier qubits, and the leading error-correction approach is the surface code. On the other hand, ion traps and neutral atoms have significantly higher fidelities but slower gate times, and therefore the requirements for error correction can be quite different. Nevertheless, most platforms aim to implement fault-tolerant Clifford gates augmented with magic state teleportation, and therefore we expect the results in this section to apply to a broad range of hardware platforms.

### 4.1. Method 1: direct gate synthesis

The most straightforward approach is a direct gate synthesis whereby the non-Clifford rotation is decomposed into a sequence of Clifford gates and typically T-gates. Here we consider approximating  $R_Z(\Delta)$  to a precision  $\epsilon$  deterministically without using ancilla qubits as in [49], which requires  $\approx 3.02 \log_2(1/\epsilon) + 1.77$  T-gates on average. Thus, in our example we estimate approximately 62 T-gates are required to synthesize each rotation gate to a precision  $\epsilon = 10^{-6}$  which adds up to a total of 2438 336 T-gates. In contrast, performing Trotterisation with  $N = 10\,000$  rounds requires 4000 000 rotation gates each with a precision  $\epsilon = 10^{-8}$  which adds up to a total of 356 000 000 T-gates. As a remark, by using approximate or non-deterministic synthesis, e.g. repeat-until-success synthesis, the costs can be reduced to  $\approx 1.03 \log_2(1/\epsilon) + 5.75$  [50–52]. In contrast, exact synthesis can also be achieved by randomly choosing from a library of short T-depth approximate rotations [40].

### 4.2. Repeat-until-success methods

The repeat-until-success approach implements rotation gates by iteratively teleporting the following resource states as

$$|\theta\rangle \equiv R_Z(\theta)|+\rangle = \frac{1}{\sqrt{2}} \left( e^{-i\theta/2}|0\rangle + e^{+i\theta/2}|1\rangle \right).$$

With probability  $1/2$  the teleportation circuit yields a measurement outcome  $+1$  which indicates that the qubit is correctly rotated as  $R_Z(\theta)|\psi\rangle$ . However, with equal chance, it yields a  $-1$  measurement outcome which indicates an inverse rotation  $R_Z(-\theta)|\psi\rangle$ . In the latter case, one needs to apply a rotation gate with twice the angle  $R_Z(2\theta)$  in order to obtain the desired  $R_Z(\theta)|\psi\rangle$  net effect. The approach is therefore repeated iteratively using resource states with angle settings  $2^k\theta$  until a  $+1$  outcome is achieved, which in general requires on average the following number of trials as

$$1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + \dots = \sum_{i=1}^{\infty} \frac{i}{2^i} = 2.$$

Using the Clifford hierarchy, which we define in appendix G, is particularly beneficial for our purposes. In particular, in the following we assume that our rotation angles are of the form  $\Delta_\ell = \pi 2^{-\ell+1}$ , thus the rotation gate  $R_Z(\Delta_\ell)$  is in the  $\ell$ th Clifford hierarchy. The significant advantage is that the repeat until success approach can terminate successfully after  $\ell-3$  unsuccessful trials given the  $\ell=3$  case is a T-gate which can be applied deterministically using T-state teleportation. Since we consider  $\Delta \equiv \Delta_{\ell_0=9}$  in the present example, the probability of termination with a T state teleportation is relatively high as  $2^{-6}$ .

We note that we can directly prepare Clifford hierarchy states  $|\Delta_\ell\rangle$  by distillation using Reed-Mueller codes [53]. This may be beneficial when  $\ell$  is sufficiently low, however for large  $\ell$ , the cost of directly distilling a resource state  $|\Delta_\ell\rangle$  to high precision will likely exceed the cost of distilling a multiple T-states that can produce  $|\Delta_\ell\rangle$  using the techniques we now introduce. Therefore, we consider alternative approaches that enable the preparation of these resource states by consuming relatively few T-states [53].

#### 4.2.1. Method 2: hamming weight phasing

Using the Hamming weight phasing method introduced in [54] and appendix A of [55], we can efficiently apply  $n$  equal-angle  $R_Z(\theta)$  rotations simultaneously to  $n$  qubits. The technique still uses  $\lceil \log_2 n + 1 \rceil$  directly synthesised, arbitrary-angle  $R_Z(\cdot)$  rotations but only uses in addition at most  $4(n-1)$  T-gates (or  $n-1$  Toffoli gates) and  $n-1$  ancilla qubits. The total required number of T-gates is therefore

$$h(n) := C_{\text{syn}}(\epsilon) \lceil \log_2 n + 1 \rceil + 4(n-1), \quad (15)$$

where  $C_{\text{syn}}(\epsilon)$  is number of T-gates required to synthesise an arbitrary-angle  $R_Z(\cdot)$  rotation to precision  $\epsilon$  and we assume that  $C_{\text{syn}} = 62$  as in Method 1.

While we could directly apply these rotations to the computational qubits, instead we apply them to  $n$  storage qubits to prepare resource states for the repeat-until-success approach – then the more storage qubits are available, the more T-gate savings this approach can yield and in the limit of infinite storage space the approach could produce  $K$  resource states using  $4(K-1)$  T-states.

We assume that in the present example we store  $n_\ell = 2^{\ell-4}$  number of  $|\Delta_\ell\rangle$  resource states, i.e. we use  $n_9 = 32$  qubits to store the states  $|\Delta_{\ell=9}\rangle$  and store only one of the  $|\Delta_{\ell=4}\rangle$ . We do not assume additional storage qubits for the  $\ell=3$  case given this is a T-state which we assume is provided natively on demand in the fault-tolerant machine. The total storage space then adds up to  $n_9 + n_8 + \dots + n_4 = \sum_{\ell=4}^9 2^{\ell-4} = 63$  qubits, whereas the total number of ancillary qubits is  $(n_9 - 1) + (n_8 - 1) + \dots + (n_4 - 1) = 57$ .

Ultimately, our aim is to power a total of  $K = 39328$  rotations using the repeat-until-success approach, thus we need to run the procedure  $R = \lceil K/n_9 \rceil$  times which in total costs  $R[h(n_9) + h(n_8) + \dots + h(n_4) + n_3]$  T-states, where  $h(n)$  is defined in equation (15) and for the 3rd level in the hierarchy we use  $n_3$  T-states. In total we thus need 1880980 T-gates and 63 storage qubits.

#### 4.2.2. Method 3: catalyst generation of resource states

A catalyst tower construction was proposed in [56], which builds on [55, 57]. The central object is a so-called catalyst circuit that we define in figure 13 and which consumes two  $|+\rangle$  states, a resource state  $|\Delta\rangle$  and a rotation gate  $R_Z(2\Delta)$ , and outputs three resource states  $|\Delta\rangle$ , thus, in effect applies two  $R_Z(\Delta)$  rotations at the cost of consuming one  $R_Z(2\Delta)$  rotation and 4 T-states. Sun *et al* [56] then stacked these catalyst circuits so that the overall circuit prepares a family of resource states  $|2^k\Delta\rangle$  by catalysis consuming only a single rotation  $R_Z(2^h\Delta)$  where  $h$  is the height of the tower.

Using our Clifford hierarchy construction with angles  $\Delta \equiv \Delta_{\ell_0}$ , the tower height can be set to  $h = \ell_0 - 3$ , given the rotation angle  $2^h\Delta_{\ell_0} = \Delta_3$  can be applied directly by T-state teleportation. Thus, the catalyst towers can be initialised by the family of resource states  $|\Delta_{\ell=4\dots\ell_0}\rangle$  (negligible initial cost) and can then continuously produce the required resource states. Thus, we prepare and store in total  $n_\ell = 2^{\ell-4}$  resource states as in the previous subsection.

In appendix G.3, we explicitly construct catalyst towers that branch out a relatively large number of catalysts at the top level to produce  $n_\ell$  resource states  $|\Delta_{\ell_0}\rangle$  and branch out fewer and fewer at the lower levels, in order to precisely output the desired exponential distribution  $n_\ell$  of the resource states  $|\Delta_\ell\rangle$  required for the repeat-until-success method. Then the total number of ancilla qubits required to produce the desired distribution  $n_\ell = 2^{\ell-4}$  is

$$\left\lceil \frac{2^{\ell_0-2} - \ell_0 + 1}{2} \right\rceil$$

and the total number of T-gates required is

$$\begin{cases} (2^{\ell_0} - 3\ell_0 + 1)/2 & \ell_0 \text{ is odd} \\ (2^{\ell_0} - 3\ell_0 + 6)/2 & \ell_0 \text{ is even} \end{cases}.$$

As in the previous section, our aim is to power a total of  $K = 39\,328$  rotations via the repeat-until-success approach. Thus, we prepare the 63 storage qubits in the resource states using  $\lceil (2^{\ell_0} - 3\ell_0 + 1)/2 \rceil = 243$  T-states in a single round, and we repeat this procedure  $\lceil K/n_9 \rceil = 1229$  times as in the previous section. Thus in total, we estimate 298 647 T-gates are required.

**Statement.** Given an arbitrarily chosen  $\ell_0$ , the expected number of rotation gates in theorem 1 is  $\nu_\infty = \text{csc}(\Delta_{\ell_0})(3 - \cos \Delta_{\ell_0})\|c\|_1 T$  and thus the expected T cost of implementing a time evolution using our catalyst approach is

$$N_{\text{Tgate}} = \begin{cases} (2^{\ell_0} - 3\ell_0 + 1)/2^{\ell_0-3} \cdot \nu_\infty & \ell_0 \text{ is odd,} \\ (2^{\ell_0} - 3\ell_0 + 6)/2^{\ell_0-3} \cdot \nu_\infty & \ell_0 \text{ is even,} \end{cases}$$

using the repeat-until-success approach with  $\sum_{\ell=4}^{\ell_0} 2^{\ell-4}$  storage qubits that store Clifford hierarchy states. It follows that the T cost can be upper bounded given  $\ell_0 > 2$  as  $N_{\text{Tgate}} \leq 8\nu_\infty$ .

## 5. Conclusion and discussion

We introduced TE-PAI to estimate observable expected values from effectively exactly time-evolved quantum states. The approach proceeds by generating a number of random circuits in classical pre-processing, whose outputs are post-processed to yield on average exact time evolution. A significant advantage of the approach is that the random circuits are built entirely of Pauli rotations  $R_\sigma(\cdot)$  using the Pauli operators  $\sigma$  in the system Hamiltonian and using only two kinds of rotation angles  $\Delta$  and  $\pi$ , which is particularly well-suited for fault-tolerant implementations. Furthermore, another significant advantage of TE-PAI is that it allows for a highly configurable trade-off between circuit depth and measurement overhead by adjusting a single parameter,  $\Delta$ , offering flexibility to fine-tune. This feature is particularly useful in NISQ and early-FTQC devices, where circuit depth and qubit coherence are the primary limitations. In the limit  $\Delta \rightarrow \pi/2$ , TE-PAI reduces to a Clifford circuit that is classically simulable—and Clifford circuits can be trivially executed in FTQCs—but requires infinite sampling overhead. In contrast, in the limit  $\Delta \rightarrow 0$  the measurement overhead vanishes at the cost of an infinitely deep quantum circuit. TE-PAI thus provides a continuous interpolation between these extremes and trades off coherent resources for incoherent repetition, i.e. by tuning  $\Delta$ , one effectively trades classical resources (sampling) against quantum resources (circuit depth) as also relevant in quantum resource theory. Furthermore, we proved that our circuits saturate the Lieb–Robinson bound in the sense that the number of gates required for simulating a total time  $T$  is directly proportional to  $T$ . By construction, TE-PAI respects the Lieb–Robinson bound on information propagation in expectation, so that quantities such as entanglement growth and state-transfer fidelity are reproduced on average as in the ideal evolution. These observations indicate the potential for TE-PAI’s randomized structure to be compatible with error-resilient strategies in quantum communication and computation.

Compared to other time-evolution algorithms, the key advantage of this approach is its ability to simulate time evolution without discretisation or algorithmic errors in the sense that finite Trotterisation error can be suppressed efficiently in classical pre-processing to arbitrarily low levels without affecting circuit depth. This is a particularly powerful feature when the aim is to simulate the evolution under time-dependent Hamiltonians. Furthermore, we require no ancillary qubits or advanced quantum

resources, only the ability to execute random circuits with rotation angles  $\Delta$  and  $\pi$ , and perform measurements. The main limitation of the approach is that its measurement overhead potentially grows exponentially unless the circuit depth is increased with growing system size. Nonetheless, an approximate version of TE-PAI can interpolate between the edge cases of exact time evolution with measurement overhead and approximate time evolution but no measurement overhead by tuning a continuous trade-off parameter  $\lambda$  from [40]. Building on this flexibility, future work could explore variance reduction techniques to further reduce sampling cost.

While TE-PAI is well-suited for NISQ applications due to its shallow depth, we develop particularly efficient fault-tolerant implementations building on the observation that the only non-Clifford gates we require are single-qubit  $R_z(\Delta)$  rotations with identical rotation angles. Our architecture prepares resource states  $R_z(\Delta)|+\rangle$  using less than 4 T-states via catalyst towers [54, 56] and applies the desired rotations via repeat until success teleportation [51].

TE-PAI can also naturally be combined with other randomised quantum protocols. First, our random circuits could be distributed via circuit cutting algorithms and phase estimation algorithms. For instance, [58] employs circuit-cutting techniques using the quasiprobabilistic decomposition framework of our method to remove entangling gates rather than to reduce overall circuit depth. Moreover, [59] proposes a randomized algorithm for phase estimation that balances gate complexity and shot noise, aiming to eliminate dependence on the sparsity of the Hamiltonian. Second, TE-PAI can be combined immediately with classical shadows by appending a layer of random measurement-basis transformation gates [16, 17]. This opens up powerful applications, such as shadow spectroscopy [20] or subspace expansion using time-evolved trial states [60]. Alternatively, one can also combine the approach with Pauli grouping techniques [18, 19]. Third, the approach can also be combined with algorithms whereby random initial states are time evolved, such as when estimating the density of states [24]. As many of these randomised protocols treat the evolution time  $T$  a random variable, TE-PAI can be used naturally to implement queries to random evolution times. Furthermore, our random circuits are composed of Pauli operators that appear in the Hamiltonian and can thus present opportunities for further optimisation through advanced compilation and transpilation tools, such as 2QAN, which was specifically designed for structurally similar Trotter circuits [35] and can significantly reduce the circuit depth by parallel execution of non-overlapping gates.

Given its simplicity, our approach is immediately deployable to a broad range of problems with applications in, e.g. quantum chemistry, materials science, combinatorial optimisation, high-energy physics, optimisation, quantum machine learning etc given time evolution is one of the most important quantum algorithmic subroutines.

## Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: [https://github.com/CKiumi/te\\_pai](https://github.com/CKiumi/te_pai).

## Acknowledgments

The authors thank Hayata Morisaki, Simon Benjamin, Gergory Boyd and Zhu Sun for helpful technical conversations. This work was supported by JST ASPIRE Japan Grant Number JPMJAP2319 and JST PRESTO, Japan, Grant Number JPMJPR25F1. B K thanks the University of Oxford for a Glasstone Research Fellowship and Lady Margaret Hall, Oxford for a Research Fellowship. The numerical modelling involved in this study made use of the Quantum Exact Simulation Toolkit (QuEST) [61] via the QuESTlink [62] frontend. B K thanks UKRI for the Future Leaders Fellowship Theory to Enable Practical Quantum Advantage (MR/Y015843/1). The authors also acknowledge funding from the EPSRC projects Robust and Reliable Quantum Computing (RoRQ, EP/W032635/1) and Software Enabling Early Quantum Advantage (SEEQA, EP/Y004655/1).

## Author contributions

Chusei Kiumi  0000-0003-3283-394X

Formal analysis (lead), Investigation (equal), Validation (lead), Visualization (lead), Writing – review & editing (equal)

Bálint Koczor  0000-0002-4319-6870

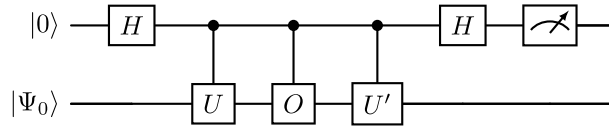
Conceptualization (lead), Investigation (equal), Methodology (lead), Writing – review & editing (equal)

## Appendix A. Comparison to prior work

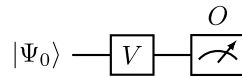
The *qDRIFT* algorithm [26] is a randomized Hamiltonian simulation method based on Hamiltonian sparsification. It was originally proposed for a time-independent Hamiltonian  $H = \sum_{j=1}^L c_j P_j$  with  $\|c\|_1 = \sum_j |c_j|$ . At each step, a term  $P_j$  is sampled with probability  $p_j = |c_j|/\|c\|_1$ , and the unitary  $e^{-i\text{sgn}(c_j)P_j\tau}$  is applied with  $\tau = T\|c\|_1/N$ , where  $T$  is the total simulation time and  $N$  is the number of steps. Averaging over the  $N$  independent random steps yields a quantum channel whose expectation matches  $e^{-iHT}$  up to an additive error  $\epsilon$ , with gate complexity  $\mathcal{O}(\|c\|_1^2 T^2/\epsilon)$ , independent of the number of Hamiltonian terms  $L$ . The rotation angle  $\tau$  is fixed across all steps of the algorithm but depends on the target accuracy  $\epsilon$  through the choice of  $N$ . In contrast, the TE-PAI algorithm with constant measurement overhead achieves gate complexity  $\mathcal{O}(\|c\|_1^2 T^2/Q)$  for a tunable parameter  $Q$ , incurring a constant overhead factor of  $\exp(Q)$ . In TE-PAI, the rotation angles are fixed at  $\pm\Delta$  and  $\pi$ , independent of the target accuracy  $\epsilon$  (though dependent on  $Q$ ). As a result, TE-PAI can achieve exact simulation with a finite gate count for any fixed  $Q$ . There are many variants of *qDRIFT*, such as for time-dependent Hamiltonians [63, 64], which can be compared to TE-PAI in a similar manner.

Next, we compare TE-PAI to a related algorithm developed in [30] which, similar to TE-PAI, simulates the time evolution by averaging over random quantum circuits. The approach decomposes the small-angle unitary rotations into linear combination of unitary matrices which are then randomly sampled to yield an unbiased estimator. In contrast, TE-PAI obtains an unbiased estimator by randomising a linear combination of superoperators in equation (8).

This leads to the following differences in practice. First, given the randomly generated quantum circuits  $U$  and  $U'$ , and target unitary observable  $O$ , the approach of [30] proceeds by controlling these quantum circuits on the state of an ancilla qubit via the Hadamard test as



In contrast, TE-PAI needs only execute a random circuit  $V$  and directly estimate observables without the need for controlling the circuits as, e.g. in the following implementation



For this reason TE-PAI is particularly well suited for near-term applications, such as NISQ or early-FTQC implementations.

Second, the approach of [30] requires a different Hadamard-test circuit for each unitary observable  $M$ . The significant advantage of TE-PAI is that it is compatible with all advanced measurement techniques and can thus be used for the simultaneous estimation of multiple observables, e.g. classical shadows, Pauli grouping, and can naturally be used to directly estimate expected values of non-unitary observables, e.g. estimating the probability of a bitstring.

Granet and Dreyer [30] bounded the measurement overhead  $\|g\|_1^\infty$  of the approach and we find it coincides with the measurement overhead of TE-PAI as

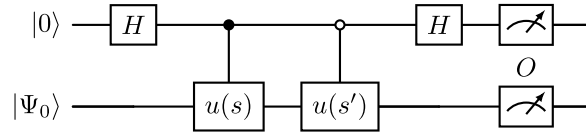
$$\exp \left[ 2\overline{\|c\|_1} T \tan \left( \frac{\Delta}{2} \right) \right].$$

Furthermore, the expected number of gates in [30] is  $2 \csc(\Delta) \overline{\|c\|_1} T$  which is approximately the same as the number of gates in TE-PAI (assuming small  $\Delta$ ) as

$$\nu_\infty = (3 - \cos \Delta) \csc(\Delta) \overline{\|c\|_1} T.$$

Furthermore, quantum simulation based on truncated Dyson series, as opposed to product formulas, is well-established in the literature [65]. A related work proposed an unbiased random circuit compiler (URCC) based on the Dyson expansion and achieved exact (on average) simulation of time-dependent Hamiltonians using an LCU expansion, continuous unbiased sampling of its terms, and a leading-order rotation to suppress variance [29]. Similar to TE-PAI, which uses fixed rotation angles  $\Delta$  and  $\pi$  (Pauli

operator), URCC applies in its leading-order branch a single Pauli rotation with a fixed angle, while its remaining-order branch consists of Pauli operations sampled from the Dyson expansion. Both URCC and TE-PAI achieve comparable finite gate counts and constant sampling overheads that are independent of the target precision. However, URCC requires complex, controlled quantum circuits as



where  $u(s)$  and  $u(s')$  are time-evolution subcircuits sampled from the Dyson series and  $O$  is the observable of interest.

In section 4, we develop fault-tolerant implementations of TE-PAI that efficiently prepare multiple copies of the resource states required for the fixed-angle rotations  $R_Z(\Delta)$ . These methods apply immediately to other techniques, such as qDRIFT and URCC, since they likewise rely on the same fixed-angle Pauli rotations discussed above. However, the number of rotation gates required by qDRIFT is  $2\|c\|_1^2 T^2/\epsilon$  for a target accuracy  $\epsilon$ . Assuming a time-independent Hamiltonian with  $\|c\|_1 = 241.3$  and  $T = 1$  (time-independent version of section 4), then for a target accuracy of  $\epsilon = 10^{-3}$ , qDRIFT would require approximately  $1.16 \times 10^8$  rotation gates, which is about three orders of magnitude more than TE-PAI in a single circuit.

Recent advances in scalable quantum simulation for constrained hardware have focused on trading quantum circuit complexity for increased classical post-processing. Similar to TE-PAI, Harrow and Lowe [58] propose *circuit cutting* techniques that simulate large circuits by partitioning them into smaller, independent fragments. Their key idea is that entangling unitaries can be replaced by ensembles of local unitaries, expressed via quasiprobabilistic decompositions  $U(\rho) = \sum_j p_j V_j^{(A)} \otimes W_j^{(B)}(\rho)$ , where  $\{p_j\}$  are quasiprobabilities and  $V_j^{(A)}, W_j^{(B)}$  are local unitaries. The reconstructed observable expectation values are given by  $\langle O \rangle = \sum_j p_j \mathbb{E}[O_j]$ , with sample variance scaling as  $\|p\|_1^2$ . To minimize this overhead, they introduce the *product extent*, and construct optimal protocols using *two-copy Hadamard tests*. One key application is the simulation of *clustered Hamiltonians* with weak inter-cluster coupling.

## Appendix B. Summary of probabilistic angle interpolation

We assume a quantum system comprising of  $N$  qubits, and consider parameterised quantum gates  $R(\theta) = e^{-i\theta G/2}$ , where  $G$  is a Pauli string as  $G \in \{\mathbb{1}, X, Y, Z\}^{\otimes N}$ . These gates are fundamental to quantum technologies given single and two-qubit rotation gates are typically engineered as Pauli gates. Here we briefly review PAI [31], which enables these gates to operate at discrete angular settings  $\Theta_k$  determined by  $B$  bits, defined as

$$\Theta_k = k\Delta, \quad \Delta = \frac{2\pi}{2^B}, \quad k \in \{0, 1, \dots, 2^B - 1\}.$$

The PAI method effectively allows for any continuous rotation angle to be achieved by overrotating from one of the discrete settings, selecting from three potential notch settings for each gate in a circuit. This approach not only ensures the desired rotation but also maintains a probability distribution centered around the same mean value as would be achieved with infinite angular resolution. The trade-off, however, is an increased number of circuit repetitions, which grows exponentially as  $e^{\nu\Delta^2/4}$  with the number of gates  $\nu$ . Nevertheless, [31] found that at a resolution of  $B = 7$  bits, the overhead is reasonable for circuits containing up to a few thousand parametrized gates, as relevant in non-error corrected machines.

We introduce the following notation for the superoperators of the aforementioned discrete-angle rotation gates as

$$\mathcal{R}_1 := R(\Theta_k), \mathcal{R}_2 := R(\Theta_{k+1}), \mathcal{R}_3 := R(\Theta_k + \pi),$$

then any overrotation  $R(\Theta_k + \theta)$  by a continuous angle  $\theta < \Delta$  can be expressed as a linear combination of the discrete gates as

$$\mathcal{R}(\Theta_k + \theta) = \gamma_1(\theta)\mathcal{R}_1 + \gamma_2(\theta)\mathcal{R}_2 + \gamma_3(\theta)\mathcal{R}_3.$$

By solving a system of trigonometric equations, Koczor *et al* [31] obtained the analytic form of the coefficients  $\gamma_l(\theta)$  as

$$\begin{aligned}\gamma_1(\theta) &= \csc\left(\frac{\Delta}{2}\right) \cos\left(\frac{\theta}{2}\right) \sin\left(\frac{\Delta}{2} - \frac{\theta}{2}\right), \\ \gamma_2(\theta) &= \csc(\Delta) \sin(\theta), \\ \gamma_3(\theta) &= -\sec\left(\frac{\Delta}{2}\right) \sin\left(\frac{\theta}{2}\right) \sin\left(\frac{\Delta}{2} - \frac{\theta}{2}\right),\end{aligned}$$

as a function of the continuous-angle  $\theta$ . We can also analytically compute the vector norm as

$$\|\gamma\|_1 = \sec\left(\frac{\Delta}{2}\right) \cos\left(\frac{\Delta}{2} - \theta\right).$$

Analogously to quasiprobability sampling methods which mitigate non-unitary error effects, PAI randomly samples the discrete rotation gates according to the above weights. In particular, we randomly choose one of the three discrete gate variants  $\{\mathcal{R}(\Theta_k), \mathcal{R}(\Theta_{k+1}), \mathcal{R}(\Theta_k + \pi)\}$  according to the probabilities  $p_l(\theta) = |\gamma_l(\theta)|/\|\gamma(\theta)\|_1$  which yields the unbiased estimator of the rotation gate as

$$\hat{\mathcal{R}}(\Theta_k + \theta) = \|\gamma(\theta)\|_1 \text{sign}[\gamma_l(\theta)] \mathcal{R}_l,$$

such that  $\mathbb{E}[\hat{\mathcal{R}}(\Theta_k + \theta)] = \mathcal{R}(\Theta_k + \theta)$ . Koczor *et al* [31] proved that PAI is optimal in the sense that the choice of the three discrete angle settings globally minimises the measurement overhead characterised by  $\|\gamma(\theta)\|_1$ .

We now briefly summarise statement 2 of [31] which is concerned with applying the PAI protocol to each continuous-angle rotation in a circuit. To simplify notations we assume a circuit  $\mathcal{U}_{\text{circ}}$  that contains no other gates than  $\nu$  parametrised ones as

$$\mathcal{U}_{\text{circ}} = \prod_{j=1}^{\nu} \mathcal{R}^{(j)}(\Theta_{k_j} + \theta_j),$$

however, it is straightforward to generalise to the case when the circuit contains other non-parametrised gates too. Here  $\mathcal{R}^{(j)}$  denotes the  $j$ th parametrised gate which is ideally set to the continuous-angle that we express as an over rotation by an angle  $\theta_j$  relative to the notch setting  $\Theta_{k_j}$ . Let us, denote the discrete rotations as

$$\begin{aligned}\mathcal{R}_1^{(j)} &:= \mathcal{R}^{(j)}(\Theta_{k_j}), \quad \mathcal{R}_2^{(j)} := \mathcal{R}^{(j)}(\Theta_{k_{j+1}}), \\ \mathcal{R}_3^{(j)} &:= \mathcal{R}^{(j)}(\Theta_{k_j} + \pi).\end{aligned}$$

At each execution of the circuit, we randomly replace a parametrised gate with the corresponding discrete gate variant, i.e, the  $j$ th parametrised gate is replaced by one of the discrete gate variants  $\mathcal{R}_l^{(j)}$ , according to the probability distribution  $p_l(\theta_j)$ . Given a circuit  $\mathcal{U}_{\text{circ}}$  of  $\nu$  parametrised gates, we choose a multi index  $\underline{l} \in 3^\nu$  according to the probability distribution  $p(\underline{l}) = |g_{\underline{l}}|/\|g\|_1$  where  $g_{\underline{l}}$

$$g_{\underline{l}} = \prod_{j=1}^{\nu} \gamma_l^{(j)}(\theta_j).$$

We obtain an unbiased estimator of the ideal circuit as

$$\hat{\mathcal{U}}_{\text{circ}} := \|g\|_1 \text{sign}(g_{\underline{l}}) \mathcal{U}_{\underline{l}},$$

where  $\|g\|_1 = \prod_{j=1}^{\nu} \|\gamma^{(j)}\|_1$ . PAI then executing the circuit variants  $\mathcal{U}_{\underline{l}}$  in which all continuously parametrised gates are replaced by the discrete ones according to the multi index  $\underline{l}$ . This yields an unbiased estimator of the ideal circuit in the sense that  $\mathbb{E}[\hat{\mathcal{U}}_{\text{circ}}] = \mathcal{U}_{\text{circ}}$ .

After performing a measurement, one multiplies the random outcome with a factor  $\|g\|_1 \text{sign}(g_{\underline{l}})$  that can have negative signs. As a consequence, the variance of the estimator is magnified which implies an increased number of circuit repetitions. Applying PAI to the estimation of an expected value results in an unbiased estimator  $\hat{o}$  of the expected value of an observable as  $\mathbb{E}[\hat{o}] = \text{Tr}[\mathcal{O}\mathcal{U}_{\text{circ}}|0\rangle\langle 0|] = o$ . The number of repetitions required to determine the expected value  $o$  to accuracy  $\epsilon$  scales as

$$N_s \leq \epsilon^{-2} \|g\|_1^2,$$

which is increased by the measurement overhead factor  $\|g\|_1^2$  compared to when having access to the ideal unitary.

### Appendix C. Classical pre-processing resources

We have finite  $N$  in classical pre-processing, and we want to bound the classical pre-processing resources. An important feature of TE-PAI is that the quantum circuit depth is independent of the Trotter number  $N$ ; increasing  $N$  only affects the classical pre-processing resources required to generate the unbiased estimators. In this subsection, we rigorously bound these classical resources.

For a fixed  $N$ , the additive Trotter error of a first-order product formula is bounded as in statement 1,

$$\epsilon_T \leq \frac{T^2}{2N} \|c\|_T^2,$$

where  $\|c\|_T$  denotes the commutator norm defined in equation (2). This error quantifies the algorithmic discrepancy between the exact unitary and the product formula at finite  $N$ . We recall that the first-order Trotter error involves a sum of commutators. Since for any pair of Pauli operators  $\|[P_i, P_j]\| \leq 2$ , we obtain

$$\begin{aligned} \|c\|_T^2 &\leq 2 \sum_{i < j} |c_i| |c_j| \\ &= \left( \sum_{j=1}^L |c_j| \right)^2 - \sum_{j=1}^L |c_j|^2 \\ &= \|c\|_1^2 - \|c\|_2^2. \end{aligned}$$

Therefore, a simple universal bound is

$$\epsilon_T \leq \frac{T^2}{2N} \|c\|_1^2,$$

which only requires the easily computable  $\ell_1$  norm of the Hamiltonian coefficients.

In practice, we require the algorithmic error  $\epsilon_T$  to be negligible compared to the statistical error  $\epsilon$  of the estimated expected value. We may demand, for some constant  $\kappa \geq 1$  which expresses how much algorithmic errors are below the statistical uncertainty, that

$$\epsilon_T \leq \kappa^{-1} \epsilon.$$

This condition ensures that algorithmic errors are several orders of magnitude smaller than the statistical uncertainty. Substituting the bound for  $\epsilon$  yields a constraint on the Trotter number,

$$N \geq \frac{T^2 \|c\|_1^2}{2\kappa^{-1} \epsilon}.$$

Thus, a sufficiently large  $N$  guarantees that TE-PAI is effectively exact to the desired precision. Therefore, the classical pre-processing cost  $\mathcal{C}$ , which corresponds to the number of classical samples required to choose an angle from three candidates, becomes

$$\mathcal{C} := NLN_s \leq \frac{1}{2} \kappa \|g\|_1^2 \|c\|_1^2 L T^2 \epsilon^{-3} = \mathcal{O}(\|g\|_1^2 \|c\|_1^2 L T^2 \epsilon^{-3}).$$

Let us illustrate this on the example of the spin-ring Hamiltonian we consider in the main text whereby each qubit is touched by seven operators, but only ten pairs per site actually anticommute, giving a maximum commutator weight of 20. Summing over all sites yields

$$\|c\|_T^2 \leq 2 \left( 4 \sum_{k=1}^n |\omega_k| + 6n \right) \leq 20n. \quad (\text{C1})$$

In our numerical demonstrations, we considered a time evolution of  $T = 1$  with  $N_s = 1000$  shots. To guarantee that the algorithmic error is within the statistical error, the above bound implies that one must take  $N \geq 140\sqrt{1000}\kappa \approx 4427\kappa$ . However, in practice the actual Trotter error is often much smaller than the above theoretical bound. Thus, in our simulation in section 3, we increased the number of Trotter steps  $N$  until the Trotter error became negligible, so that the results matched the exact curve within statistical error bounds. In this way, we found that  $N = 1000$  is sufficient in practice.

## Appendix D. Proof of expected number of gates

### D.1. Mean value

**Proof.** We now detail our derivation of the expected number  $\mathbb{E}(\nu)$  of gates in TE-PAI circuits. As PAI replaces each continuously parametrised gate in a circuit with one of three discrete gate variants, and because one of those three options is the identity operation which does not increase the number of gates, the number of gates is a Bernoulli distributed random variable. More specifically, the operational dynamics of the gates are defined such that at position  $k$  and  $j$ , exactly one gate is added to the circuit with probability  $1 - p_1(|\theta_{kj}|)$  (either  $\pm\Delta$  rotation or  $\pi$  rotation), while the identity operation is selected with probability  $p_1(|\theta_{kj}|)$ . Given that we implement this selection process across  $N$  time steps and  $L$  different gate types, we effectively conduct  $NL$  independent trials. This allows us to compute the expected total number of gates as

$$\mathbb{E}(\nu) = \sum_{j=1}^N \sum_{k=1}^L (1 - p_1(|\theta_{kj}|)).$$

We express the probability  $p_1$  in terms of the gate parameters  $|\theta_{kj}|$ , where:

$$p_1(|\theta_{kj}|) = \frac{\sin(\Delta - |\theta_{kj}|) + \sin(\Delta) - \sin(|\theta_{kj}|)}{2(\sin(\Delta - |\theta_{kj}|) + \sin(|\theta_{kj}|))}.$$

Expanding these for large  $N$  we obtain the series for  $p_1$  as

$$\begin{aligned} p_1(|\theta_{kj}|) &= 1 - \frac{1}{2}(3 - \cos\Delta) \csc(\Delta) |\theta_{kj}| \\ &\quad + \left( \sec^2\left(\frac{\Delta}{2}\right) - \frac{1}{4} \right) |\theta_{kj}|^2 + O(|\theta_{kj}|^3) \\ &= 1 - (3 - \cos\Delta) \csc(\Delta) |c_k(t_j)| \frac{T}{N} + \mathcal{D}_{kj}, \end{aligned}$$

where in the second equation we introduced the error term

$$\begin{aligned} \mathcal{D}_{kj} &:= \left( \sec^2\left(\frac{\Delta}{2}\right) - \frac{1}{4} \right) |\theta_{kj}|^2 + O(|\theta_{kj}|^3) \\ &\leq 8|c_k(t_j)|^2 \frac{T^2}{N^2} + O(N^{-3}) \end{aligned}$$

that we bound using  $(\sec^2(\frac{\Delta}{2}) - \frac{1}{4}) \leq 2$  given that  $\Delta \leq \pi/2$ , and substitute in the rotation angles  $|\theta_{kj}| = 2|c_k(t_j)| \frac{T}{N}$ .

With this approximation, the expected number of gates can be expanded for large  $N$  as

$$\begin{aligned} \mathbb{E}(\nu) &= \sum_{j=1}^N \sum_{k=1}^L \left( (3 - \cos\Delta) \csc(\Delta) |c_k(t_j)| \frac{T}{N} - \mathcal{D}_{kj} \right) \\ &= \csc(\Delta) (3 - \cos\Delta) \sum_{k=1}^L \sum_{j=1}^N \left( |c_k(t_j)| \frac{T}{N} \right) - \mathcal{E}, \end{aligned} \quad (\text{D1})$$

where we denoted the sum of individual error terms as

$$\mathcal{E} := \sum_{k=1}^L \sum_{j=1}^N \mathcal{D}_{kj} \leq 8 \sum_{k=1}^L \sum_{j=1}^N |c_k(t_j)|^2 \frac{T^2}{N^2} + O(N^{-2}).$$

We can relate the summation in equation (D1) to the Riemann integration in equation (4) as

$$\int_0^T \sum_{k=1}^L |c_k(t)| dt = \sum_{k=1}^L \sum_{j=1}^N \left( |c_k(t_j)| \frac{T}{N} \right) + \mathcal{Q}, \quad (\text{D2})$$

where we can bound the error term as  $|\mathcal{Q}| \leq V_{tot} T/N$ . In the present work we assume that the time-dependent Hamiltonian coefficients  $c_k(t)$  are absolutely continuous functions of time and therefore have bounded variation with  $V_{tot} = \sum_k V(c_k) < \infty$  being a finite constant. In case if  $c_k(t)$  is differentiable, then  $V(c_k)$  is upper bounded by the absolute largest value of the derivative function  $c_k'(t)$ , but our results also

apply to functions that are not necessarily differentiable everywhere. This follows from results of [66, 67] which proved that the error term in the Riemann summation  $\int_a^b f(x)dx = \sum_{k=1}^N f(k\Delta_N)\Delta_N + Q$  is upper bounded as  $|Q| \leq V\Delta_N$  using the notation  $\Delta_N = (b - a)/N$  and  $V$  is the absolute largest variation  $V$  of the function.

Therefore, we finally obtain the explicit expression as

$$\begin{aligned} \mathbb{E}(\nu) &= \csc(\Delta)(3 - \cos\Delta) \sum_{k=1}^L \sum_{j=1}^N \left( |c_k(t_j)| \frac{T}{N} \right) - \mathcal{E} \\ &= \csc(\Delta)(3 - \cos\Delta) \sum_{k=1}^L \int_0^T |c_k(t)| dt + \mathcal{Q}' - \mathcal{E} \\ &= \csc(\Delta)(3 - \cos\Delta) \overline{\|c\|_1} T + \mathcal{Q}' - \mathcal{E}. \end{aligned}$$

Here we can bound the two error terms using the notation  $\mathcal{Q}' = \csc(\Delta)(3 - \cos\Delta)\mathcal{Q}$  as

$$\begin{aligned} |\mathcal{Q}' - \mathcal{E}| &\leq |\mathcal{Q}'| + |\mathcal{E}| \\ &\leq \frac{3 \csc(\Delta) TV_{tot}}{N} + 8 \sum_{k=1}^L \sum_{j=1}^N |c_k(t_j)|^2 \frac{T^2}{N^2} + O(N^{-2}) \\ &= \frac{3 \csc(\Delta) TV_{tot}}{N} + 8 \overline{\|c\|_2^2} \frac{T^2}{N} + O(N^{-2}). \end{aligned}$$

In the last equation, we again replaced the Riemann summation with an integral to obtain the mean of the square of the  $\ell_2$  norm of the coefficient vector  $\overline{\|c\|_2^2}$  up to an error  $O(N^{-2})$ .

Finally, we conclude our proof as

$$\mathbb{E}(\nu) = \csc(\Delta)(3 - \cos\Delta) \overline{\|c\|_1} T + O(N^{-1}). \tag{D3}$$

In figure 9 we numerically plot  $\mathbb{E}(\nu)$  for increasing  $N$  for the time-dependent Hamiltonian introduced in section 3 and confirm that indeed it converges to the above expression. □

### D.2. Minimal number of gates

We can also compute the minimum of the number of gates over  $\Delta$ . From theorem 2, we already knew that the sampling cost is upper bounded by  $\exp Q$  if we set  $\Delta = 2 \arctan\left(\frac{Q}{2\overline{\|c\|_1}T}\right)$ . Thus, Substituting this angle yields

$$\lim_{N \rightarrow \infty} \mathbb{E}(\nu) = \frac{2\left(\overline{\|c\|_1}T\right)^2}{Q} + Q.$$

Noting that we used the following relations to  $(3 - \cos\Delta) \csc(\Delta)$ .

$$\sin(2 \arctan(x)) = \frac{2x}{1+x^2}, \quad \cos(2 \arctan(x)) = \frac{1-x^2}{1+x^2}.$$

From the assessment of the arithmetic-geometric mean, the expression  $\frac{2(\overline{\|c\|_1}T)^2}{Q} + Q$  takes its minimum value of  $\overline{\|c\|_1}T2\sqrt{2}$  when  $Q = \overline{\|c\|_1}T\sqrt{2}$  i.e.  $\Delta = 2 \tan\left(\frac{1}{\sqrt{2}}\right)$ .

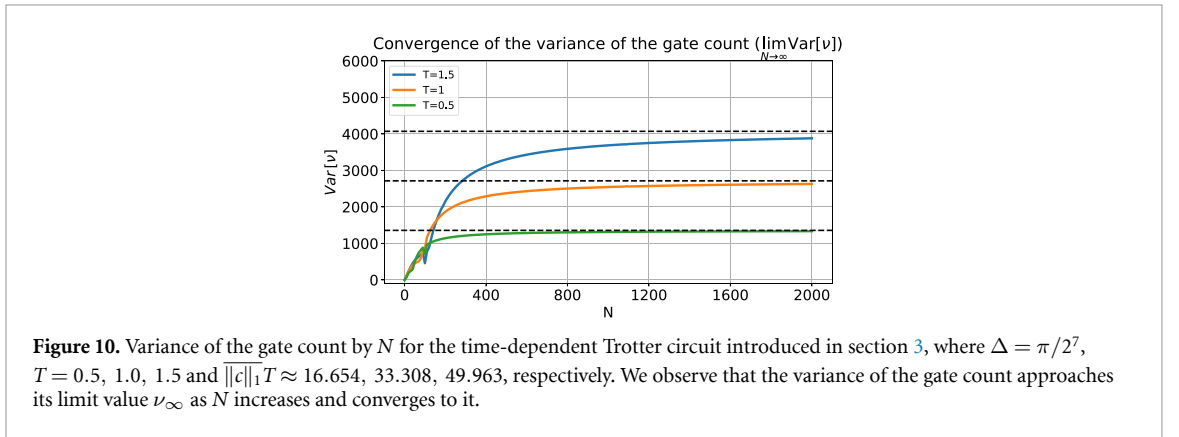
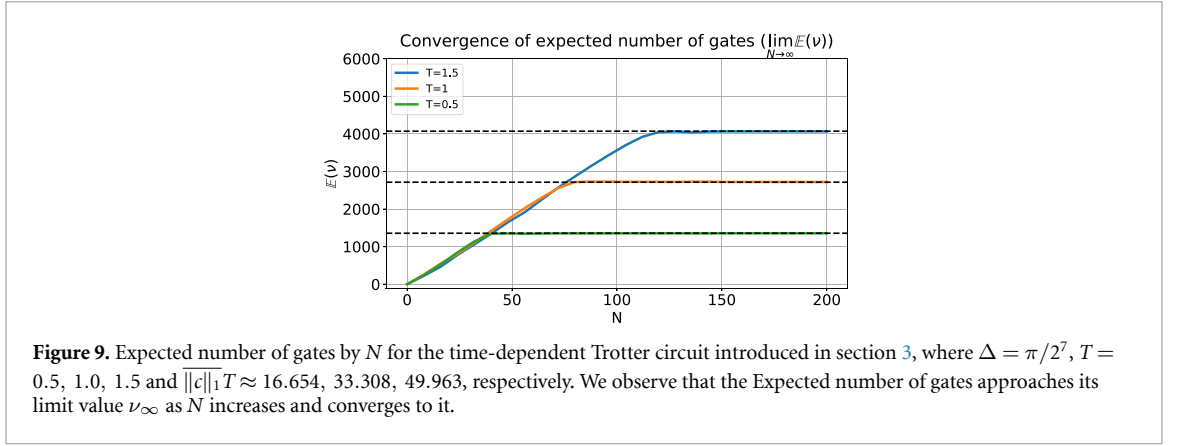
### D.3. Variance of the number of gates

The variance of the number of gates can be computed analytically given each position follows a Bernoulli distribution for which the variance is  $p(1 - p)$  and thus

$$\text{Var}[\nu] = \sum_{j=1}^N \sum_{k=1}^L p_1(|\theta_{kj}|) (1 - p_1(|\theta_{kj}|)).$$

Here the total variance of the number of gates is upper bounded by the expectation value of the number of gates as

$$\text{Var}[\nu] \leq \sum_{j=1}^N \sum_{k=1}^L (1 - p_1(|\theta_{kj}|)) = \mathbb{E}[\nu].$$



For large  $N$ , the variance can be simplified as follows:

$$\begin{aligned} \text{Var}[\nu] &= \sum_{j=1}^N \sum_{k=1}^L \text{csc}(\Delta) (3 - \cos \Delta) |c_k(t_j)| \frac{T}{N} + O(N^{-2}) \\ &= \text{csc}(\Delta) (3 - \cos \Delta) \overline{\|c\|_1} T + O(N^{-1}). \end{aligned}$$

Above we do not detail all steps as we used the argument as in equation (D2) to replace the summation with an integral, and to bound the error term as  $O(N^{-1})$ . This leads to:

$$\lim_{N \rightarrow \infty} \text{Var}[\nu] = \text{csc}(\Delta) (3 - \cos \Delta) \overline{\|c\|_1} T = \lim_{N \rightarrow \infty} \mathbb{E}[\nu].$$

As such, since we consider large  $N$ , By Lyapunov Central Limit Theorem, we can approximate the distribution of gate numbers well by the normal distribution:

$$\mathcal{N}(\mathbb{E}[\nu], \mathbb{E}[\nu]).$$

We numerically plot the variance of the number of gates in figure 10 for the time-dependent Hamiltonian introduced in section 3, respectively, and confirm convergence to the above expressions.

## Appendix E. Proof of measurement overhead

**Proof.** In this section we provide a detailed derivation of the measurement overhead in PAI applied within Trotter circuits for simulating time-dependent Hamiltonian systems. As we detailed above, the measurement overhead of PAI is bounded by the following expression as

$$\|g\|_1 = \prod_{j=1}^N \prod_{k=1}^L \|\gamma_k(|\theta_{kj}\rangle)\|_1,$$

which quantifies the cumulative measurement overhead of the circuit by considering the overhead introduced by individual rotation gates.

The measurement overhead for each gate is given by:

$$\begin{aligned} \|\gamma(|\theta_{kj}|)\|_1 &= \sec\left(\frac{\Delta}{2}\right) \cos\left(\frac{\Delta}{2} - |\theta_{kj}|\right) \\ &= \cos(|\theta_{kj}|) + \sin(|\theta_{kj}|) \tan\left(\frac{\Delta}{2}\right). \end{aligned}$$

We again consider the limit of large  $N$  and thus small angles  $\theta_k(t)$ , and obtain the series expansion as

$$\|\gamma(|\theta_{kj}|)\|_1 = 1 + \tan\left(\frac{\Delta}{2}\right) |\theta_{kj}| + \mathcal{D}_{kj},$$

where also substitute in our expression for the angles  $\theta_{kj}$  and introduce the error term  $\mathcal{D}_{kj} \in O(N^{-2})$  as

$$\mathcal{D}_{kj} = -|\theta_{kj}|^2/2 + O(\theta_{kj}^3) = -2|c_k(t_j)|^2 \frac{T^2}{N^2} + O(N^{-3}).$$

The total measurement cost  $\|g\|_1$  can thus be evaluated by expanding each term in the product as

$$\|g\|_1 = \prod_{j=1}^N \prod_{k=1}^L \left[ 1 + 2 \tan\left(\frac{\Delta}{2}\right) |c_k(t_j)| \frac{T}{N} + \mathcal{D}_{kj} \right].$$

We now take the logarithm of  $\|g\|_1$ , which allows us to convert the product into a sum as

$$\begin{aligned} \log \|g\|_1 &= \sum_{j=1}^N \sum_{k=1}^L \log \left[ 1 + 2 \tan\left(\frac{\Delta}{2}\right) |c_k(t_j)| \frac{T}{N} + \mathcal{D}_{kj} \right] \\ &= \sum_{j=1}^N \sum_{k=1}^L \left[ 2 \tan\left(\frac{\Delta}{2}\right) |c_k(t_j)| \frac{T}{N} + \mathcal{D}_{kj} \right] \\ &= 2 \tan\left(\frac{\Delta}{2}\right) \sum_{j=1}^N \sum_{k=1}^L |c_k(t_j)| \frac{T}{N} + \mathcal{E}, \end{aligned} \tag{E1}$$

where in the second equation we used the expansion  $\log(1+x+\epsilon) = \log(1+x) + \epsilon/(1+x) + O(\epsilon^2)$  and  $\log(1+x) = x - x^2/2 + O(x^3)$  for small  $\epsilon = \mathcal{D}_{kj} \in O(N^{-2})$  and  $x = 2 \tan(\frac{\Delta}{2}) |c_k(t_j)| \frac{T}{N} \in O(N^{-1})$ , respectively. We introduce the error term as

$$\mathcal{E} := \sum_{j=1}^N \sum_{k=1}^L \mathcal{D}_{kj}. \tag{E2}$$

We can again replace the summation in equation (E1) with Riemann integration as introduced in equation (D2) as

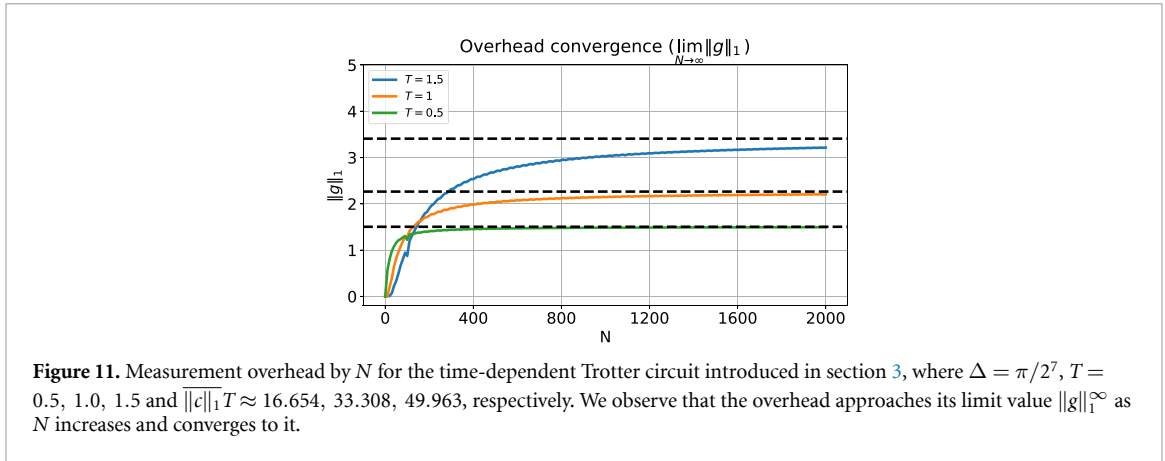
$$\sum_{j=1}^N \sum_{k=1}^L |c_k(t_j)| \frac{T}{N} = \sum_{k=1}^L \int_0^T |c_k(t)| dt + \mathcal{Q}, \tag{E3}$$

up to the error term  $|\mathcal{Q}| \leq V_{tot}T/N$ . This yields

$$\log \|g\|_1 = 2 \tan\left(\frac{\Delta}{2}\right) \sum_{k=1}^L \int_0^T |c_k(t)| dt + \mathcal{Q}' + \mathcal{E},$$

where we use the notation  $\mathcal{Q}' = 2 \tan(\frac{\Delta}{2}) \mathcal{Q}$ . We finally obtain

$$\begin{aligned} \|g\|_1 &= \exp \left[ 2 \tan\left(\frac{\Delta}{2}\right) \overline{\|c\|}_1 T + \mathcal{Q}' + \mathcal{E} \right] \\ &= \exp \left[ 2 \tan\left(\frac{\Delta}{2}\right) \overline{\|c\|}_1 T \right] \exp[\mathcal{Q}' + \mathcal{E}], \end{aligned}$$



**Figure 11.** Measurement overhead by  $N$  for the time-dependent Trotter circuit introduced in section 3, where  $\Delta = \pi/2^7$ ,  $T = 0.5, 1.0, 1.5$  and  $\|c\|_1 T \approx 16.654, 33.308, 49.963$ , respectively. We observe that the overhead approaches its limit value  $\|g\|_1^\infty$  as  $N$  increases and converges to it.

where we used the notation  $\overline{\|c\|_1} = \frac{1}{T} \int_0^T \sum_{k=1}^L |c_k(t)| dt$  from equation (4), we conclude our proof by bounding the error terms as

$$\begin{aligned} |Q' + \mathcal{E}| &\leq |Q'| + |\mathcal{E}| \\ &\leq 2 \tan\left(\frac{\Delta}{2}\right) \frac{TV_{tot}}{N} + \sum_{j=1}^N \sum_{k=1}^L |\mathcal{D}_{kj}| \\ &= 2 \tan\left(\frac{\Delta}{2}\right) \frac{TV_{tot}}{N} + 2\overline{\|c\|_1^2} \frac{T^2}{N} + O(N^{-2}) \end{aligned}$$

which yields the expression

$$\|g\|_1 = \exp\left[2 \tan\left(\frac{\Delta}{2}\right) \overline{\|c\|_1} T\right] + O(N^{-1}). \quad (\text{E4})$$

□

In figure 11 we numerically plot  $\|g\|_1$  for increasing  $N$  for the time-dependent Hamiltonian introduced in section 3 and confirm that indeed it converges to the above expression.

## Appendix F. Proof of remark 1

In remark 1, we introduced a control trade-off parameter  $Q$  that manages the trade-off between the circuit depth and the number of shots required. The expression for the number of gates  $\nu_\infty$  is given by:

$$\nu_\infty = \frac{2\left(\overline{\|c\|_1} T\right)^2}{Q} + Q.$$

Here, we have two competing factors as the term  $\frac{2\left(\overline{\|c\|_1} T\right)^2}{Q}$  decreases and the second term increases as  $Q$  increases. We can rewrite the expression as

$$\nu_\infty = \frac{2\left(\overline{\|c\|_1} T\right)^2}{Q} + \frac{Q^2}{Q} = \frac{2\left(\overline{\|c\|_1} T\right)^2 + Q^2}{Q}.$$

The lower bound of  $\nu_\infty$  in theorem 1 is attained when  $Q = \overline{\|c\|_1} T\sqrt{2}$ , and now, we substitute this bound  $Q = \overline{\|c\|_1} T\sqrt{2}$  into the expression as

$$\nu_\infty \leq \frac{2\left(\overline{\|c\|_1} T\right)^2 + \left(\overline{\|c\|_1} T\sqrt{2}\right)^2}{Q} = \frac{4\left(\overline{\|c\|_1} T\right)^2}{Q}$$

gives the upper bound, holds for  $Q \leq \overline{\|c\|_1} T\sqrt{2}$ .

## Appendix G. Details of resource estimation

### G.1. Mapping Pauli rotations to Z rotations

In FTQC, Clifford gates are considered relatively cheap to implement as natural operations in stabiliser codes. A Clifford operator on a quantum system described by  $n$  qubits is a unitary operator  $U$  such that for any Pauli operator  $P$ , the operator  $UPU^\dagger$  is also a Pauli operator. Formally, the Clifford group  $\mathcal{C}$  is defined as:

$$\mathcal{C} := \{U \mid UPU^\dagger \in \mathcal{P}, \forall P \in \mathcal{P}\},$$

where  $\mathcal{P}$  denotes the Pauli group. Clifford gates are pivotal in quantum error correction schemes as they are efficiently implementable and largely error-free compared to non-Clifford gates.

Non-Clifford gates, such as T-gates, on the other hand, are more challenging and costly to implement due to the need for non-trivial additional measures, such as magic state distillation. Thus, minimizing the number of non-Clifford gates, such as continuous-angle rotations, is crucial for minimising the resources requirements in early-FTQC.

In our TE-PAI approach, we require only discrete Pauli rotations  $R_\sigma(\theta)$  of the form  $e^{i\theta\sigma/2}$ , where  $\sigma$  represents Pauli strings. These rotations can be efficiently mapped to single-qubit  $Z$  rotations, interleaved with Clifford gates. More precisely, any Pauli rotation  $R_\sigma(\theta)$  can be expressed as

$$R_\sigma(\theta) = UR_Z(\theta)U^\dagger,$$

where  $U$  is a sequence of Clifford gates determined by the specific Pauli operator  $\sigma$  and the rotation angle  $\theta$ . This formulation allows us to focus on efficiently implementing the  $R_Z(\pm\Delta)$  rotation gate which is then the only non-Clifford resource we require.

### G.2. Clifford hierarchy

The Clifford hierarchy, denoted as  $\mathcal{C}_\ell$ , is defined recursively, beginning with the Pauli group  $\mathcal{P}$  at the first level as  $\mathcal{C}_1 := \mathcal{P}$  and for  $\ell > 1$ , the higher levels of the hierarchy are defined as

$$\mathcal{C}_\ell := \{U \mid UPU^\dagger \in \mathcal{C}_{\ell-1}, \forall P \in \mathcal{P}\}.$$

This recursive definition means that the unitary  $U$  at level  $\ell$  conjugates elements of the Pauli group  $\mathcal{P}$  to operators in  $\mathcal{C}_{\ell-1}$ . Thus,  $\mathcal{C}_2 = \mathcal{C}$  becomes a Clifford group.

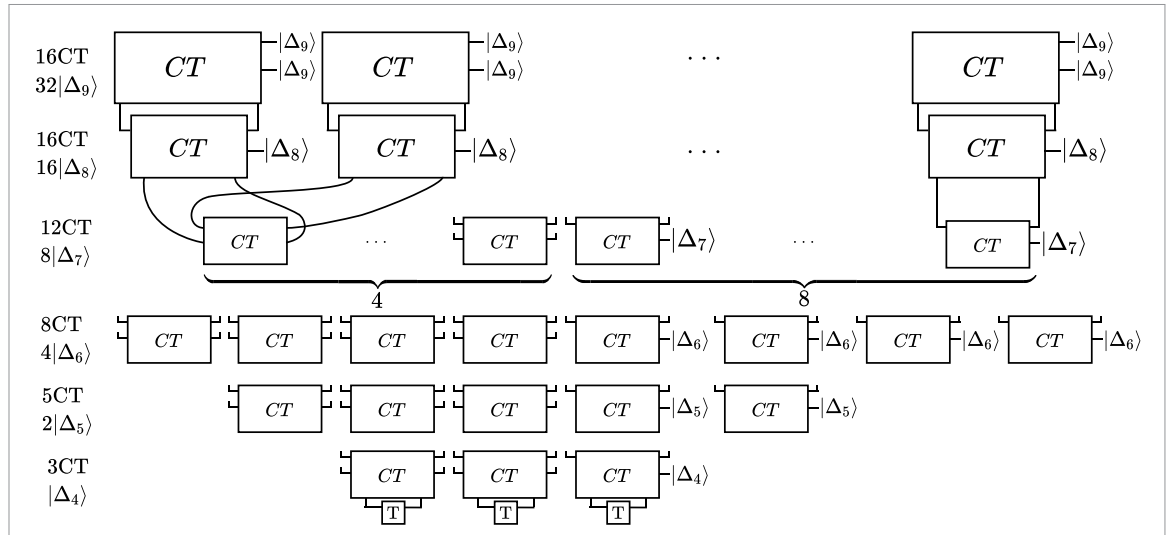
### G.3. Catalyst towers for Clifford hierarchy rotations

Here, we explain how we construct a catalyst tower to generate resource states in Clifford hierarchy for repeat-until-success method. The construction of the catalyst tower is based on [56]. In figure 12, we give the constructive example for  $\ell_0 = 9$ , but our construction is straightforwardly generalises to higher  $\ell_0$ .

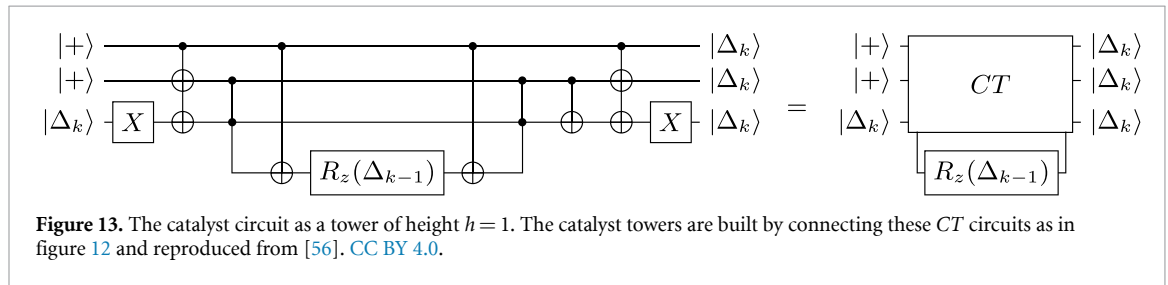
The white boxes in figure 12 indicate the catalyst circuits that were introduced in [56] and which we denote as  $CT$  and we define them explicitly in figure 13. While [56] concatenated these circuits to yield a catalyst tower that outputs an approximately equal number of  $|\Delta_\ell\rangle$  resource states, our construction in figure 12 outputs resource states  $|\Delta_\ell\rangle$  according to an exponential distribution as required for the repeat-until-success implementation of the rotation gate  $R_Z(\Delta_\ell)$ . Specifically, our catalyst tower outputs  $2^{\ell-4}$  resource states of  $|\Delta_\ell\rangle$  for  $\ell = 4, 5, \dots, \ell_0$ , and we explicitly demonstrate the case of  $\ell_0 = 9$ .  $|\Delta_3\rangle$  is a T-state, and we assume that T-states are natively produced by the fault-tolerant quantum hardware, e.g. via magic state distillation.

The tower is constructed as follows:

- **Top Layer:** In the top layer, each  $CT$  circuit generates two  $|\Delta_{\ell_0}\rangle$  states. Therefore, we need  $2^{\ell_0-5}$   $CT$  circuits at the top layer to generate all required  $2^{\ell_0-4}$  of  $|\Delta_{\ell_0}\rangle$  states.
- **Second Layer:** Each  $CT$  circuit in the second layer is connected to exactly one  $CT$  circuit in the top layer. The total number of  $CT$  circuits in the second layer is  $2^{\ell_0-5}$ , which generates  $2^{\ell_0-5}$  number of  $|\Delta_{\ell_0-1}\rangle$  states.
- **Third Layer:** In the third layer, we aim to generate  $2^{\ell_0-6}$  number of  $|\Delta_{\ell_0-2}\rangle$  states, so  $2^{\ell_0-6}$   $CT$  circuits are required. To ensure we provide enough  $|\Delta_{\ell_0-2}\rangle$  states to the  $CT$  circuits in the second layer, we use an additional  $(2^{\ell_0-5} - 2^{\ell_0-6})/2 = 2^{\ell_0-7}$  extra  $CT$  circuits, which do not generate additional resource states but are used solely to support the generation of  $|\Delta_{\ell_0-2}\rangle$  states for the second layer. Thus in total, we need  $3 \times 2^{\ell_0-7}$   $CT$  circuits.



**Figure 12.** The catalyst tower to generate resource states in Clifford hierarchy of  $\ell_0 = 9$  which follows an exponential distribution, i.e. generating  $2^{\ell-4}$  resource states of  $|\Delta_\ell\rangle$  for  $\ell = 4, 5, \dots, \ell_0 = 9$ . In the top layer, since each CT circuit generates two  $|\Delta_9\rangle$  states, we require  $2^{\ell-5}$  CT circuits. In the second layer, each CT circuit is connected to exactly one CT circuit from the first layer, resulting in 16 CT circuits to generate 16  $|\Delta_8\rangle$  states. For the third layer, since we need to generate 8  $|\Delta_7\rangle$  states, we need 8 CT circuits. Additionally, to provide  $|\Delta_7\rangle$  to the remaining CT circuits in the second layer, we require  $(16 - 8)/2 = 4$  extra CT circuits, which do not generate additional  $|\Delta_7\rangle$  states. The same approach applies to the following layers. In the final layer, since  $|\Delta_4\rangle$  corresponds to the T-state, we can directly apply three T-gates for this layer. Therefore, in total, we need 60 CT circuits and 3 T-gates, resulting in a total cost of 243 T-gates and 60 ancilla qubits.



**Figure 13.** The catalyst circuit as a tower of height  $h = 1$ . The catalyst towers are built by connecting these CT circuits as in figure 12 and reproduced from [56]. CC BY 4.0.

- **Remaining Layers:** The same process continues for subsequent layers, where we progressively halve the number of  $|\Delta_\ell\rangle$  states generated at each layer. At every step, additional CT circuits are used to provide resource states to the above layer, following the same pattern.
- **Final Layer:** In the final layer, the resource state  $|\Delta_3\rangle$  corresponds to the T-gate, so we can directly apply T-gates to complete the process.

Based on the above, by using mathematical induction, we can calculate that  $(\ell_0 - \ell + 1)2^{\ell-5}$  CT circuits are required at each layer  $\ell$ . Note that if  $\ell_0$  is an even number, this expression for the final layer  $\ell = 4$  is not an integer. Thus, it will require one extra T-gate and CT circuit that generate second  $|\Delta_4\rangle$ . Therefore, the total number of CT circuits required is calculated as follows:

$$\left\lceil \sum_{\ell=4}^{\ell_0} (\ell_0 - \ell + 1) 2^{\ell-5} \right\rceil = \left\lceil \frac{2^{\ell_0-2} - \ell_0 + 1}{2} \right\rceil.$$

The number of ancillary qubits is equal to the number of CT circuits. Additionally,  $\lceil (\ell_0 - 3)/2 \rceil$  T-gates are applied to the final layer. Since each CT circuit requires 4 T gates, the total T cost of the entire process becomes  $(2^{\ell_0} - 3\ell_0 + 1)/2$  for odd  $\ell_0$  and  $(2^{\ell_0} - 3\ell_0 + 6)/2$  for even  $\ell_0$ .

## References

[1] Feynman R P 1982 Simulating physics with computers *Int. J. Theor. Phys.* **21** 467–88  
 [2] Lloyd S 1996 Universal quantum simulators *Science* **273** 1073–8  
 [3] Suzuki M 1990 Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations *Phys. Lett. A* **146** 319  
 [4] Suzuki M 1991 General theory of fractal path integrals with applications to many-body theories and statistical physics *J. Math. Phys.* **32** 400–7

- [5] Reiher M, Wiebe N, Svore K M, Wecker D and Troyer M 2017 Elucidating reaction mechanisms on quantum computers *Proc. Natl Acad. Sci.* **114** 7555–60
- [6] McArdle S, Endo S, Aspuru-Guzik A, Benjamin S C and Yuan X 2020 Quantum computational chemistry *Rev. Mod. Phys.* **92** 015003
- [7] Bauer B, Bravyi S, Motta M and Chan G K-L 2020 Quantum algorithms for quantum chemistry and quantum materials science *Chem. Rev.* **120** 12685–717
- [8] Su Y, Berry D W, Wiebe N, Rubin N and Babbush R 2021 Fault-tolerant quantum simulations of chemistry in first quantization *PRX Quantum* **2** 040332
- [9] Childs A M and Wiebe N 2012 Hamiltonian simulation using linear combinations of unitary operations *Quantum Inf. Comput.* **12** 901–24
- [10] Berry D W, Childs A M, Cleve R, Kothari R and Somma R D 2014 Exponential improvement in precision for simulating sparse Hamiltonians *Proc. 46th Annual ACM Symp. on Theory of Computing (STOC '14)* (ACM) (<https://doi.org/10.1145/2591796.2591854>)
- [11] Berry D W, Childs A M, Cleve R, Kothari R and Somma R D 2015 Simulating Hamiltonian dynamics with a truncated Taylor series *Phys. Rev. Lett.* **114** 090502
- [12] Low G H and Chuang I L 2017 Optimal Hamiltonian simulation by quantum signal processing *Phys. Rev. Lett.* **118** 010501
- [13] Low G H and Chuang I L 2019 Hamiltonian simulation by qubitization *Quantum* **3** 163
- [14] Childs A M and Berry D W 2012 Black-box Hamiltonian simulation and unitary implementation *Quantum Inf. Comput.* **12** 29–62
- [15] Childs A M 2009 On the relationship between continuous- and discrete-time quantum walk *Commun. Math. Phys.* **294** 581–603
- [16] Inane H, Steinberg J, Cai Z, Nguyen H C and Koczor B 2024 Quantum error mitigated classical shadows *PRX Quantum* **5** 010324
- [17] Huang H-Y, Kueng R and Preskill J 2020 Predicting many properties of a quantum system from very few measurements *Nat. Phys.* **16** 1050–7
- [18] Crawford O, van Straaten B, Wang D, Parks T, Campbell E and Brierley S 2021 Efficient quantum measurement of Pauli operators in the presence of finite sampling error *Quantum* **5** 385
- [19] Jena A, Genin S and Mosca M 2019 Pauli partitioning with respect to gate sets (arXiv:1907.07859)
- [20] Chan H H S, Meister R, Goh M L and Koczor B 2025 Algorithmic shadow spectroscopy *PRX Quantum* **6** 010352
- [21] Wang G, França D S, Zhang R, Zhu S and Johnson P D 2023 Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision *Quantum* **7** 1167
- [22] Lin L and Tong Y 2022 Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers *PRX Quantum* **3** 010318
- [23] Ding Z and Lin L 2023 Simultaneous estimation of multiple eigenvalues with short-depth quantum circuit on early fault-tolerant quantum computers *Quantum* **7** 1136
- [24] Goh M L and Koczor B 2024 Direct estimation of the density of states for fermionic systems (arXiv:2407.03414)
- [25] Yang Y, Lu B-N and Li Y 2021 Accelerated quantum monte carlo with mitigated error on noisy quantum computer *PRX Quantum* **2** 040361
- [26] Campbell E 2019 Random compiler for fast Hamiltonian simulation *Phys. Rev. Lett.* **123** 070503
- [27] Ouyang Y, White D R and Campbell E T 2020 Compilation by stochastic Hamiltonian sparsification *Quantum* **4** 235
- [28] Chen C-F, Huang H-Y, Kueng R and Tropp J A 2021 Concentration for random product formulas *PRX Quantum* **2** 040305
- [29] Zhang X-M, Huo Z, Liu K, Li Y and Yuan X 2022 Unbiased random circuit compiler for time-dependent Hamiltonian simulation (arXiv:2212.09445)
- [30] Granet E and Dreyer H 2024 Continuous Hamiltonian dynamics on digital quantum computers without discretization error *npj Quantum Inf.* **10** 82
- [31] Koczor B, Morton J J L and Benjamin S C 2024 Probabilistic interpolation of quantum rotation angles *Phys. Rev. Lett.* **132** 130602
- [32] Lieb E H and Robinson D W 1972 The finite group velocity of quantum spin systems *Commun. Math. Phys.* **28** 251–7
- [33] Haah J, Hastings M B, Kothari R and Low G H 2021 Quantum algorithm for simulating real time evolution of lattice hamiltonians *SIAM J. Comput.* **52** FOCS18
- [34] Kiumi C 2025 Python implementation of TE-PAI (available at: [https://github.com/CKiumi/te\\_pai](https://github.com/CKiumi/te_pai))
- [35] Lao L and Browne D E 2022 2QAN: a quantum compiler for 2-local qubit Hamiltonian simulation algorithms *Proc. 49th Annual Int. Symp. on Computer Architecture (ISCA '22)* (Association for Computing Machinery) pp 351–65
- [36] Childs A M, Su Y, Tran M C, Wiebe N and Zhu S 2021 Theory of trotter error with commutator scaling *Phys. Rev. X* **11** 011020
- [37] An D, Fang D and Lin L 2022 Time-dependent Hamiltonian simulation of highly oscillatory dynamics and superconvergence for Schrödinger equation *Quantum* **6** 690
- [38] Koch C P *et al* 2022 Quantum optimal control in quantum technologies. Strategic report on current status, visions and goals for research in Europe *EPJ Quantum Technol.* **9** 19
- [39] Blanes S, Casas F, Oteo J and Ros J 2009 The magnus expansion and some of its applications *Phys. Rep.* **470** 151
- [40] Koczor B 2024 Sparse probabilistic synthesis of quantum operations *PRX Quantum* **5** 040352
- [41] Lee J, Berry D W, Gidney C, Huggins W J, McClean J R, Wiebe N and Babbush R 2021 Even more efficient quantum computations of chemistry through tensor hypercontraction *PRX Quantum* **2** 030305
- [42] Labib F, Clader B D, Stamatopoulos N and Zeng W J 2024 Quantum amplitude estimation from classical signal processing (arXiv:2405.14697)
- [43] Gilyén A, Su Y, Low G H and Wiebe N 2019 Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics *Proc. 51st Annual ACM SIGACT Symp. on Theory of Computing (STOC '19)* (ACM) pp 193–204
- [44] Luitz D J, Laflorencie N and Alet F 2015 Many-body localization edge in the random-field heisenberg chain *Phys. Rev. B* **91** 081103
- [45] Childs A M, Maslov D, Nam Y, Ross N J and Su Y 2018 Toward the first quantum simulation with quantum speedup *Proc. Natl Acad. Sci.* **115** 9456–61
- [46] Cai Z, Babbush R, Benjamin S C, Endo S, Huggins W J, Li Y, McClean J R and O'Brien T E 2023 Quantum error mitigation *Rev. Mod. Phys.* **95** 045005
- [47] Dalzell A M, Hunter-Jones N and Brandão F G 2024 Random quantum circuits transform local noise into global white noise *Commun. Math. Phys.* **405** 78
- [48] Foldager J and Koczor B 2023 Can shallow quantum circuits scramble local noise into global white noise? *J. Phys. A: Math. Theor.* **57** 015306
- [49] Ross N J and Selinger P 2016 Optimal ancilla-free Clifford+T approximation of z-rotations *Quantum Inf. Comput.* **16** 901–53

- [50] Bocharov A, Roetteler M and Svore K M 2015 Efficient synthesis of probabilistic quantum circuits with fallback *Phys. Rev. A* **91** 052317
- [51] Bocharov A, Roetteler M and Svore K M 2015 Efficient synthesis of universal repeat-until-success quantum circuits *Phys. Rev. Lett.* **114** 080502
- [52] Kliuchnikov V, Lauter K, Minko R, Paetznick A and Petit C 2023 Shorter quantum circuits via single-qubit gate approximation *Quantum* **7** 1208
- [53] Landahl A J and Cesare C 2013 Complex instruction set computing architecture for performing accurate quantum  $Z$  rotations with less magic (arXiv:1302.3240)
- [54] Gidney C 2018 Halving the cost of quantum addition *Quantum* **2** 74
- [55] Kivlichan I D *et al* 2020 Improved fault-tolerant quantum simulation of condensed-phase correlated electrons via Trotterization *Quantum* **4** 296
- [56] Sun Z, Boyd G, Cai Z, Jnane H, Koczor B, Meister R, Minko R, Pring B, Benjamin S C and Stamatopoulos N 2024 Low depth phase oracle using a parallel piecewise circuit (arXiv:2409.04587)
- [57] Gidney C and Fowler A G 2019 Efficient magic state factories with a catalyzed CCZ to 2T transformation *Quantum* **3** 135
- [58] Harrow A W and Lowe A 2025 Optimal quantum circuit cuts with application to clustered Hamiltonian simulation *PRX Quantum* **6** 010316
- [59] Wan K, Berta M and Campbell E T 2022 Randomized quantum algorithm for statistical phase estimation *Phys. Rev. Lett.* **129** 030503
- [60] Boyd G, Koczor B and Cai Z 2024 High-dimensional subspace expansion using classical shadows (arXiv:2406.11533)
- [61] Jones T, Brown A, Bush I and Benjamin S C 2019 Quest and high performance simulation of quantum computers *Sci. Rep.* **9** 10736
- [62] Jones T and Benjamin S 2020 QuESTlink—mathematica embiggened by a hardware-optimised quantum emulator *Quantum Sci. Technol.* **5** 034012
- [63] Berry D W, Childs A M, Su Y, Wang X and Wiebe N 2020 Time-dependent Hamiltonian simulation with  $L^1$ -norm scaling *Quantum* **4** 254
- [64] Cao Y, Jin S and Liu N 2024 A unifying framework for quantum simulation algorithms for time-dependent Hamiltonian dynamics (arXiv:2411.03180)
- [65] Kieferová M, Scherer A and Berry D W 2019 Simulating the dynamics of time-dependent hamiltonians with a truncated Dyson series *Phys. Rev. A* **99** 042314
- [66] Owens L 2014 Exploring the rate of convergence of approximations to the Riemann integral *Preprint*
- [67] Tasaki H 2009 Convergence rates of approximate sums of Riemann integrals *J. Approx. Theory* **161** 477