

Averaging gate approximation error and performance of Unitary Coupled Cluster ansatz in Pre-FTQC Era

Kohdai Kuroiwa^{1,2,3,4} and Yuya O. Nakagawa⁴

¹Institute for Quantum Computing, University of Waterloo, Ontario, Canada, N2L 3G1

²Department of Combinatorics and Optimization, University of Waterloo

³Perimeter Institute for Theoretical Physics, Ontario, Canada, N2L 2Y5

⁴QunaSys Inc., Aqua Hakusan Building 9F, 1-13-7 Hakusan, Bunkyo, Tokyo 113-0001, Japan

Fault-tolerant quantum computation (FTQC) is essential to implement quantum algorithms in a noise-resilient way, and thus to enjoy advantages of quantum computers even with presence of noise. In FTQC, a quantum circuit is decomposed into universal gates that can be fault-tolerantly implemented, for example, Clifford+ T gates. Here, T gate is usually regarded as an essential resource for quantum computation because its action cannot be simulated efficiently on classical computers and it is experimentally difficult to implement fault-tolerantly. Practically, it is highly likely that only a limited number of T gates are available in the near future. In this *Pre-FTQC* era, due to the constraint on available resources, it is vital to precisely estimate the decomposition error of a whole circuit. In this paper, we propose that the Clifford+ T decomposition error for a given quantum circuit containing a large number of quantum gates can be modeled as the depolarizing noise by averaging the decomposition error for each quantum gate in the circuit, and our model provides more accurate error estimation than the naive estimation. We exemplify this by taking unitary coupled-cluster (UCC) ansatz used in the applications of quantum computers to quantum chemistry as an example. We theoretically evaluate the approximation error of UCC ansatz when decomposed into Clifford+ T gates, and the numerical simulation for a wide variety of molecules verified that our model well explains the total decomposition error of the ansatz. Our results enable the precise and efficient usage of quantum resources in the early-stage applications of quantum computers and fuel further research towards what quantum computation can achieve in the upcoming future.

1 Introduction

Quantum computers are expected to exhibit superior performance over classical computers. For example, Shor's factoring algorithm [1] shows an exponential speed-up over any known classical algorithm, and Grover's searching algorithm [2] achieves a quadratic speed-up. In

Kohdai Kuroiwa: kkuroiwa@uwaterloo.ca

Yuya O. Nakagawa: nakagawa@qunasys.com

recent years, owing to advancements in technology, the experimental realization of quantum computers has made steady progress, and quantum supremacy has even been demonstrated using these devices [3–8]. Noisy intermediate-scale quantum (NISQ) devices [9] are currently largely investigated as a promising pioneer of quantum computational technologies. However, in the future, we aim to perform complex quantum algorithms, which show substantial quantum speed-ups, such as the phase estimation algorithm [10, 11] and factoring algorithm [1]. It is inevitable that circuits required in these algorithms become highly deep, and noise occurring during the computation has significant effects without error correction. Thus, eventually, we need fault-tolerant quantum computation (FTQC) using quantum error correction to maximize the benefit of quantum computers.

Fault-tolerant implementation of quantum circuits with a family of error correction codes [12, 13] requires an approximation of a given quantum circuit using universal gates that can be fault-tolerantly implemented with the error correction codes. Solovay-Kitaev (SK) theorem [14, 15] gives an algorithm that generates decomposition of a given quantum gate into a desired universal gate set that is closed under inversion. For example, SK theorem is applicable to the Clifford+ T universal gate set, and this set is often considered as a gate set for FTQC due to its fault-tolerant implementability [16–21]. Hence, in this work, we focus on the circuit approximation with the Clifford+ T universal gate set. In this context, T gate is considered as a crucial resource in FTQC because its fault-tolerant implementation, associating the magic state distillation [22], is the most costly part in terms of both time and space overhead with current practical error-correcting codes. Let us also note that Gottesman-Knill Theorem [23, 24] guarantees that quantum circuits consisting only of Clifford gates are efficiently simulatable by classical computers.

Practically, it is highly likely that we do not have a sufficient resource, *e.g.*, T gates and clean qubits, in the near future due to its difficulty of experimental implementation. In such an early stage of FTQC, *Pre-FTQC era*, the precise estimation of the error arising from the decomposition of a given quantum circuit into Clifford+ T gates must be crucial to taking full advantage of available quantum computers. There has been growing attention to Pre-FTQC (or early-FTQC) era recently, including algorithms with less auxiliary qubits [25] and improved noise robustness [26], error reduction using error mitigation methods [27], techniques shortening the circuit depth [28], a method for constructing a state preparation circuit [29], and variational algorithms with limited number of T gates [30]. Recently, a model for this early stage of FTQC was developed by Ref. [31]. In Ref. [32], an architecture for Pre-FTQC, where we partially perform error correction, was also proposed.

Nevertheless, precise error estimation of the Clifford+ T decomposition of quantum circuits has not yet been thoroughly investigated. Typically, the accuracy of such decompositions based on the SK theorem is assessed by summing the errors of individual gate decompositions via the triangle inequality [13]. However, this naive evaluation may not provide a tight bound when the quantum circuit is large and the decomposition errors are non-uniform across gates. Practical resource estimation in Pre-FTQC era may well demand more precise error estimation, which requires another way to evaluate the error. The best accuracy we can achieve using the available number of T gates reveals how many T gates will be necessary to realize desired computation, so it will greatly help the hardware development of quantum computers as well as the practical use of them.

2 Contribution

In this work, we propose a method for accurate estimation of the error caused by the Clifford+ T decomposition for a given quantum circuit, when the circuit contains a suf-

ficiently large number of gates. We argue that the Clifford+ T approximation error of a quantum circuit can be treated as a single-qubit depolarizing channel by averaging the approximation error of each gate in the circuit, which largely simplifies the analysis of the approximation error of the whole circuit. While the naive estimation by the triangle inequality yields the total circuit approximation error scaling linearly in given gate approximation error ϵ , the error estimation given by our averaging method results is in a quadratic form of it, $O(\epsilon^2)$, as shown in Eq. (11) in Sec. 4. In this sense, our method succeeds in reducing the estimated error of Clifford+ T circuit approximation.

We illustrate our method in Unitary Coupled Cluster (UCC) ansatz [33], which has been recognized as a good parametrized state for quantum chemistry problems in near-term quantum computing. This ansatz can also be applied in the Pre-FTQC era, for example, for preparation of an initial state of quantum phase estimation [10, 11], so evaluation of the performance of the UCC ansatz within our framework of Pre-FTQC should be practically beneficial. We analyze Clifford+ T approximation error of the UCC ansatz using the averaging method discussed above, and show that Clifford+ T approximation error in this case can be expressed by a simple formula. Furthermore, we perform numerical simulation to verify this analysis for a wide variety of molecules. To see relations between the number of T gates and the precision of the circuit decomposition, we develop a Clifford+ T -gate decomposition algorithm for single-qubit rotation gates with the fixed maximum number of T gates by slightly modifying the algorithm given by Ross and Selinger [34]. With the modified decomposition algorithm, we see that our averaging method yields precise estimation of the circuit approximation error when a given UCC ansatz is large. We lastly investigate how many T -gates are needed to realize sufficiently accurate UCC ansatz and predict the number of T gates for a classically intractable large molecule.

Thus, we open up a practical perspective to implement a quantum circuit in the Pre-FTQC era. Our analysis and estimation will be useful to assess how many T gates are needed to realize a given target quantum state in the Pre-FTQC era. In this sense, we newly pave a way to develop an application of Pre-FTQC from a practical point of view, in which the number of available T -gates is restricted.

Now we briefly compare our results to related work. First, Campbell [35] demonstrated that the probabilistic application of quantum gates with an error of $O(\epsilon)$, sampled from an appropriate distribution, can approximate a target unitary with an improved error of $O(\epsilon^2)$. Our findings differ from this result in that we do not rely on probabilistic implementations. Instead, the sequence of quantum gates used to approximate the target unitary is fixed. In our numerical examples, the accumulation of errors for many quantum gates in the circuit may invoke a self-averaging effect, resulting in a quadratic error suppression of $O(\epsilon^2)$. Second, several methods such as twirling [36] or dynamical decoupling [37] are often introduced to mitigate quantum noise by averaging the effect of quantum (noisy) gates. Our approach differs from these methods fundamentally in its setup. We do not explicitly exploit random unitary gates, and thus the gate sequence remains deterministic throughout.

The rest of this paper is organized as follows. In Sec. 3, we summarize notation used in this paper. In Sec. 4, we show our method of averaging unitary Clifford+ T approximation error into the depolarizing noise. In Sec. 5, we apply our analysis to the UCC ansatz. In Sec. 5.1, we show the concrete Clifford+ T approximation error estimation for the UCC ansatz using our averaging method. In Sec. 5.2, we introduce our algorithm for generating the best Clifford+ T -gate decomposition within some given number of T gates, and compare this algorithm with the one by Ref. [34] on which it is based. In Sec. 5.3, we perform numerical experiments to verify our algorithm and error analysis using the UCC ansatz for

various molecules in quantum chemistry. Finally, in Sec. 6, we summarize our results and discuss future directions.

3 Notation

In this section, we introduce the notation in this paper. For any $\theta \in \mathbb{R}$ and any vector $\mathbf{n} = (n_x, n_y, n_z) \in \mathbb{R}^3$ satisfying $n_x^2 + n_y^2 + n_z^2 = 1$, we define the single-qubit rotation gate $R_{\mathbf{n}}(\theta)$ along axis $\mathbf{n}$ with angle θ as

$$R_{\mathbf{n}}(\theta) = \exp\left(-i\frac{\theta}{2}(n_x\sigma_x + n_y\sigma_y + n_z\sigma_z)\right), \quad (1)$$

where σ_x , σ_y , and σ_z are the Pauli operators defined as

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2)$$

in the computational basis. Using Eq. (2), we can rewrite $R_{\mathbf{n}}(\theta)$ in Eq. (1) as

$$R_{\mathbf{n}}(\theta) = \begin{pmatrix} \cos\frac{\theta}{2} - in_z\sin\frac{\theta}{2} & -i(n_x - in_y)\sin\frac{\theta}{2} \\ -i(n_x + in_y)\sin\frac{\theta}{2} & \cos\frac{\theta}{2} + in_z\sin\frac{\theta}{2} \end{pmatrix}$$

in the computational basis. In particular, when $\mathbf{n} = (0, 0, 1)$, $R_{\mathbf{n}}(\theta)$ represents the single-qubit z -rotation gate with angle θ , and we use $R_z(\theta)$ to denote this.

For a linear operator A , the trace distance of A is defined as

$$\|A\|_1 := \text{Tr} \left[\sqrt{A^\dagger A} \right],$$

and the operator norm of A is defined as

$$\|A\| := \max_{|\phi\rangle} \sqrt{\langle\phi| A^\dagger A |\phi\rangle}, \quad (3)$$

where the maximum is taken over all pure states $|\phi\rangle$.

4 Averaging of approximation error of rotation gate on a large parametrized circuit

In this section, we propose an approach to analyze the effects of gate approximation error, considering a parametrized circuit that is supposed to be still used in the Pre-FTQC era.

Typically, such a parametrized circuit, *e.g.*, the UCC ansatz circuit (See Appendix B for more details), can be expressed by Clifford gates and single-qubit z -rotation gates. Hence, when we consider Clifford+ T approximation of this circuit, we will consider approximation of z -rotation gates contained in this circuit. Conventionally, the total Clifford+ T approximation error is roughly evaluated by using the triangle inequality (See also Remark 1 in Sec. 5.1), where Clifford+ T approximation error is linearly accumulated. However, this analysis might overestimate the total error, which would be a problem when a circuit size is large.

Here, we propose an alternative method to evaluate the total Clifford+ T approximation error, which we call *averaging*. We analyze an averaged effect of Clifford+ T approximation error, or the error coming from the decomposition of a z -rotation gate in a given

parametrized circuit. By using this averaging analysis, we estimate the Clifford+ T decomposition error of the circuit. Following our analysis, one may model the effective behavior of Clifford+ T approximation error by a depolarizing channel. The averaged effect we derive is expected to be valid when there are a large number of gates to be decomposed so that Clifford+ T approximation error of each gate in the circuit will be well-averaged over the whole circuit.

Note that while our main focus is Clifford+ T approximation for a parametrized quantum circuit, the averaging method we propose below applies to any decomposition of z -rotation.

Now, let $R_z(\theta)$ be the desired single-qubit z -rotation gate with angle θ , and suppose that we have a unitary gate U as an approximation of $R_z(\theta)$ to accuracy $\epsilon > 0$, that is,

$$\|U - R_z(\theta)\| \leq \epsilon. \quad (4)$$

When we try to apply the gate $R_z(\theta)$, we actually apply the noisy gate $U = (UR_z^\dagger(\theta))R_z(\theta)$. Thus, $UR_z^\dagger(\theta)$ is considered to represent the unitary noise induced by the decomposition. Since any single-qubit unitary is equivalent to some single-qubit rotation gate $R_{\mathbf{n}}(\varphi)$ along axis $\mathbf{n} = (n_x, n_y, n_z)$ with angle φ up to some global phase $e^{i\alpha}$, we can express the unitary noise as $UR_z^\dagger(\theta) = e^{i\alpha}R_{\mathbf{n}}(\varphi)$ using real numbers $\alpha, \varphi \in (-\pi, \pi]$ and a normalized real vector $\mathbf{n}$. Since the operator norm is unitary-invariant, from Eq. (4), we have

$$\|I - e^{i\alpha}R_{\mathbf{n}}(\varphi)\| \leq \epsilon. \quad (5)$$

To satisfy Eq. (5), all the singular values of $I - e^{i\alpha}R_{\mathbf{n}}(\varphi)$ must be upper bounded by ϵ . Consider the following positive matrix:

$$\begin{aligned} & (I - e^{i\alpha}R_{\mathbf{n}}(\varphi))(I - e^{i\alpha}R_{\mathbf{n}}(\varphi))^\dagger \\ &= 2I - e^{i\alpha}R_{\mathbf{n}}(\varphi) - e^{-i\alpha}R_{\mathbf{n}}^\dagger(\varphi) \\ &= 2 \begin{pmatrix} 1 - \cos \alpha \cos \frac{\varphi}{2} - n_z \sin \alpha \sin \frac{\varphi}{2} & -(n_x - in_y) \sin \alpha \sin \frac{\varphi}{2} \\ -(n_x + in_y) \sin \alpha \sin \frac{\varphi}{2} & 1 - \cos \alpha \cos \frac{\varphi}{2} + n_z \sin \alpha \sin \frac{\varphi}{2} \end{pmatrix}. \end{aligned}$$

The eigenvalues of this matrix are

$$2 \left(1 - \cos \left(\frac{\varphi}{2} \pm \alpha \right) \right).$$

Therefore, for Eq. (5) to be satisfied, the following must hold:

$$1 - \cos \left(\frac{\varphi}{2} + \alpha \right) \leq \frac{\epsilon^2}{2}, \quad (6)$$

$$1 - \cos \left(\frac{\varphi}{2} - \alpha \right) \leq \frac{\epsilon^2}{2}. \quad (7)$$

Define $\delta := \cos^{-1}(1 - \epsilon^2/2)$. By this definition, when $\epsilon \approx 0$, we also have $\delta \approx 0$. Since $\alpha, \varphi \in (-\pi, \pi]$, these conditions, Eq. (6) and (7), are equivalent to

$$-2\delta - 2\alpha \leq \varphi \leq 2\delta - 2\alpha, \quad (8)$$

$$-2\delta + 2\alpha \leq \varphi \leq 2\delta + 2\alpha, \quad (9)$$

that is, α and φ must satisfy

$$\left| \frac{\varphi}{2} \pm \alpha \right| \leq \delta.$$

Additionally, $\mathbf{n}$ is constrained by the condition $|\mathbf{n}|^2 = n_x^2 + n_y^2 + n_z^2 = 1$.

Now, we analyze the average effect of the error caused by the approximation of a single qubit z -rotation. We show the effect of the noise $UR_z^\dagger(\theta) = e^{i\alpha}R_{\mathbf{n}}(\varphi)$ averaged over the parameters is indeed equivalent to the depolarizing noise. Assuming that the effect of noise is completely random over the circuit, we take the Haar measure for the average. Hence, the average effect of $R_{\mathbf{n}}(\varphi)$ on a single qubit operator A can be expressed as

$$\int_{|\frac{\varphi}{2} \pm \alpha| \leq \delta} \sin^2 \frac{\varphi}{2} d\varphi d\alpha \int_{|\mathbf{n}|=1} dn_x dn_y dn_z R_{\mathbf{n}}(\varphi) A R_{\mathbf{n}}^\dagger(\varphi), \quad (10)$$

where the term $\sin^2 \frac{\varphi}{2}$ comes from the Haar measure. Note that the normalization is not yet taken into account at this point.

First, we evaluate the integral with respect to $\mathbf{n}$ in Eq. (10). The integral can be written as

$$\begin{aligned} & \int_{|\mathbf{n}|=1} dn_x dn_y dn_z R_{\mathbf{n}}(\varphi) A R_{\mathbf{n}}^\dagger(\varphi) \\ &= \int_{|\mathbf{n}|=1} dn_x dn_y dn_z \left(\cos^2 \frac{\varphi}{2} A + \sin^2 \frac{\varphi}{2} \left(n_x^2 \sigma_x A \sigma_x + n_y^2 \sigma_y A \sigma_y + n_z^2 \sigma_z A \sigma_z \right) \right. \\ & \quad + i \cos \frac{\varphi}{2} \sin \frac{\varphi}{2} (n_x (A \sigma_x - \sigma_x A) + n_y (A \sigma_y - \sigma_y A) + n_z (A \sigma_z - \sigma_z A)) \\ & \quad \left. + \sin^2 \frac{\varphi}{2} (n_x n_y (\sigma_x A \sigma_y + \sigma_y A \sigma_x) + n_y n_z (\sigma_y A \sigma_z + \sigma_z A \sigma_y) + n_z n_x (\sigma_z A \sigma_x + \sigma_x A \sigma_z)) \right). \end{aligned}$$

Since

$$\begin{aligned} & \int_{|\mathbf{n}|=1} dn_x dn_y dn_z n_i = 0 \\ & \int_{|\mathbf{n}|=1} dn_x dn_y dn_z n_j n_k = 0 \end{aligned}$$

for all $i = x, y, z$ and $j, k = x, y, z$ with $j \neq k$, the third and the fourth terms in the integral will be zero. Thus, using the fact that

$$\begin{aligned} & \int_{|\mathbf{n}|=1} dn_x dn_y dn_z = 4\pi, \\ & \int_{|\mathbf{n}|=1} dn_x dn_y dn_z n_i^2 = \frac{4\pi}{3} \end{aligned}$$

for $i = x, y, z$, we have

$$\int_{|\mathbf{n}|=1} dn_x dn_y dn_z R_{\mathbf{n}}(\varphi) A R_{\mathbf{n}}^\dagger(\varphi) = 4\pi \left(\cos^2 \frac{\varphi}{2} A + \frac{1}{3} \sin^2 \frac{\varphi}{2} (\sigma_x A \sigma_x + \sigma_y A \sigma_y + \sigma_z A \sigma_z) \right).$$

The normalization constant for this integral is

$$\int_{|\mathbf{n}|=1} dn_x dn_y dn_z = 4\pi.$$

Hence, the resulting average operator is

$$\cos^2 \frac{\varphi}{2} A + \frac{1}{3} \sin^2 \frac{\varphi}{2} (\sigma_x A \sigma_x + \sigma_y A \sigma_y + \sigma_z A \sigma_z).$$

Now, we evaluate the integral with respect to α and δ . Observe that

$$\begin{aligned}\int_{|\frac{\varphi}{2} \pm \alpha| \leq \delta} \sin^2 \frac{\varphi}{2} \cos^2 \frac{\varphi}{2} d\varphi d\alpha &= \frac{8\delta^2 + \cos 4\delta - 1}{16}, \\ \int_{|\frac{\varphi}{2} \pm \alpha| \leq \delta} \sin^4 \frac{\varphi}{2} d\varphi d\alpha &= \frac{24\delta^2 + 16 \cos 2\delta - \cos 4\delta - 15}{16}.\end{aligned}$$

Note that the normalization constant for this integral is given by

$$\int_{|\frac{\varphi}{2} \pm \alpha| \leq \delta} \sin^2 \frac{\varphi}{2} d\varphi d\alpha = 2\delta^2 + \cos 2\delta - 1.$$

Thus, after the normalization, the resulting operator is

$$\frac{8\delta^2 + \cos 4\delta - 1}{16(2\delta^2 + \cos 2\delta - 1)} A + \frac{1}{3} \frac{24\delta^2 + 16 \cos 2\delta - \cos 4\delta - 15}{16(2\delta^2 + \cos 2\delta - 1)} (\sigma_x A \sigma_x + \sigma_y A \sigma_y + \sigma_z A \sigma_z).$$

Since we have Taylor expansions

$$\begin{aligned}\frac{8\delta^2 + \cos 4\delta - 1}{16(2\delta^2 + \cos 2\delta - 1)} &= 1 - \frac{2\delta^2}{5} + O(\delta^4), \\ \frac{24\delta^2 + 16 \cos 2\delta - \cos 4\delta - 15}{16(2\delta^2 + \cos 2\delta - 1)} &= \frac{2\delta^2}{5} + O(\delta^4),\end{aligned}$$

when δ is sufficiently small, the resulting average operator can approximately be written as

$$\left(1 - \frac{2\delta^2}{5}\right) A + \frac{2\delta^2}{15} (\sigma_x A \sigma_x + \sigma_y A \sigma_y + \sigma_z A \sigma_z) = \left(1 - \frac{8\delta^2}{15}\right) A + \frac{8\delta^2}{15} \text{Tr}(A) \frac{I}{2}.$$

In summary, the average effect on A can be written as

$$A \mapsto \left(1 - \frac{8\delta^2}{15}\right) A + \frac{8\delta^2}{15} \text{Tr}(A) \frac{I}{2};$$

that is, the effect is equivalent to the single-qubit depolarizing noise with error probability

$$p := \frac{8\delta^2}{15} = O(\epsilon^2), \tag{11}$$

where we used the Taylor series

$$\cos^{-1}(1 - x) = \sqrt{2x} + O(x^{3/2}).$$

Hence, when averaged, the unitary gate-decomposition error of size ϵ is represented as depolarizing error whose magnitude is $O(\epsilon^2)$. Note that the naive estimation by the triangle inequality gives $O(\epsilon)$ [13]. We will exemplify this with the UCC ansatz in Eq. (15) in Sec. 5.1. Hence, we obtained a completely new error formula by averaging. We also justify our analysis by considering a "variance" operator with respect to averaging in Appendix A. As noted at the beginning of this section, we did not restrict the approximation to a specific form, so the analysis conducted in this section applies to any approximation method of z -rotation gates.

We believe that the averaged error derived in this section appropriately models the effects of Clifford+ T approximation error of a large quantum circuit, which consists of many

single-qubit rotations. Clifford+ T approximation error originated from the Clifford+ T decomposition of each single-qubit rotation gate is absolutely unitary by definition. At the same time, when we look at the effects of all of these unitary errors over the large circuit, it is expected that their effects look to be averaged. Therefore, from our analysis here, we anticipate that Clifford+ T approximation error of a large quantum circuit with many single-qubit rotations is well-modeled by the single-qubit depolarizing noise. In Sec. 5, we numerically examine this using Unitary Coupled Cluster (UCC) ansatz.

5 Application to Unitary Coupled Cluster (UCC) ansatz

In this section, we investigate Clifford+ T approximation error of the UCC ansatz both theoretically and numerically, as an illustration of our results. In Sec. 5.1, we analyze how the averaged Clifford+ T approximation error on the UCC ansatz behaves based on the theoretical argument of Sec. 4. In Sec. 5.2, we present our algorithm for generating the best Clifford+ T -gate decomposition within some given number of T gates, which is motivated by the algorithm given in Ref. [34]. In Sec. 5.3, we numerically investigate Clifford+ T approximation error of the UCC ansatz for various molecular systems in quantum chemistry when the number of T gates is restricted. The averaged Clifford+ T approximation error well describes the numerical results and we also estimate the actual number of the T gates required to represent a quantum state with sufficient accuracy for the first time.

In the Pre-FQTC era, while we cannot perform powerful quantum algorithms that require tremendous amounts of resources, we may still wish to prepare specific quantum states that are impossible to realize on near-term quantum devices without error correction, *e.g.*, parametrized quantum states realized by a relatively deep ansatz circuit. In quantum chemistry, although the quantum phase estimation [10, 11] is expected to compute the ground state energy of a given molecular Hamiltonian efficiently, its realization will require about $O(10^{10}) \sim O(10^{15})$ T gates for molecules of practical interest [38–43]. It is then natural to consider using early FTQC with a smaller number of T -gates, such as the variational quantum eigensolver [44] with an ansatz with a deep quantum circuit, at some point in time. In the context of near-term quantum computation, the UCC ansatz [33], which is based on the coupled cluster theory [45], has been considered as a nice parametrized state (ansatz state) to represent a ground state of a molecular Hamiltonian. While the application of the UCC ansatz has been primarily discussed for near-term quantum devices so far, we believe that this ansatz will rather be beneficial in the Pre-FTQC era because its circuit tends to be pretty deep in general and typically beyond the capability of near-term quantum devices¹. We also note that the UCC ansatz with some fixed parameters can serve as an initial state of quantum phase estimation, which may overcome the problem of the poor overlap between the initial state and the exact ground state of the molecular Hamiltonian [46]. We summarize the definition and properties of the UCC ansatz in Appendix B (see also Ref. [33] for more details).

5.1 Averaged effects of Clifford+ T approximation error on UCC ansatz

We evaluate Clifford+ T approximation error of the UCC ansatz with large number of qubits by using the average effects of Clifford+ T approximation error derived in Sec. 4. Observing that quantum gates in the UCC ansatz can be expressed as a composition of H gates, S gates, CNOT gates, and parametrized single-qubit z -rotation gates, (see

¹The UCC singles and doubles ansatz, which is used in the numerical simulation below, requires $O(N^4)$ gates for N molecular orbitals in a naive implementation.

Appendix B for the case of UCC singles and doubles ansatz), we know that Clifford+ T approximation error arises only when we approximate the z -rotation by the Clifford+ T gate decomposition. Suppose that we have a UCC ansatz circuit with L z -rotation gates. Then, the UCC ansatz circuit with the averaged Clifford+ T approximation error can be modeled as a quantum operation

$$\mathcal{N}_{\text{UCC}} := \mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_2 \circ \mathcal{U}_2 \circ \mathcal{E}_1 \circ \mathcal{U}_1,$$

where $\mathcal{E}_i$ is the single-qubit depolarizing channel acting on the single-qubit system to which the i th z -rotation is applied, and $\mathcal{U}_i$ is a unitary operation composed of Clifford gates and only one z -rotation gate, for all $1 \leq i \leq L$. Denoting the initial input to this ansatz ρ_0 , the total effects of Clifford+ T approximation error of the circuit can be evaluated by

$$\epsilon_{\text{tot}} := \|\mathcal{N}_{\text{UCC}}(\rho_0) - \mathcal{U}_{\text{UCC}}(\rho_0)\|_1,$$

where

$$\mathcal{U}_{\text{UCC}} := \mathcal{U}_L \circ \cdots \circ \mathcal{U}_2 \circ \mathcal{U}_1$$

is the given UCC ansatz circuit and $\|\cdot\|_1$ is the trace norm. Observe that

$$\begin{aligned} \mathcal{N}_{\text{UCC}}(\rho_0) &= \mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_2 \circ \mathcal{U}_2 \circ \mathcal{E}_1 \circ \mathcal{U}_1(\rho_0) \\ &= \mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_2 \circ \mathcal{U}_2 \circ ((1-p)\mathcal{U}_1(\rho_0) + p\Delta_1 \circ \mathcal{U}_1(\rho_0)) \\ &= (1-p)\mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_2 \circ \mathcal{U}_2 \circ \mathcal{U}_1(\rho_0) \\ &\quad + p\mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_2 \circ \Delta_1 \circ \mathcal{U}_1(\rho_0) \\ &= (1-p)\mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_2 \circ \mathcal{U}_2 \circ \mathcal{U}_1(\rho_0) + p\rho_1 \\ &= (1-p)\mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_3 \circ \mathcal{U}_3 \circ ((1-p)\mathcal{U}_2 \circ \mathcal{U}_1(\rho_0) + p\Delta_2 \circ \mathcal{U}_2 \circ \mathcal{U}_1(\rho_0)) + p\rho_1 \\ &= (1-p)^2\mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_3 \circ \mathcal{U}_3 \circ \mathcal{U}_2 \circ \mathcal{U}_1(\rho_0) \\ &\quad + p(1-p)\mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_3 \circ \Delta_2 \circ \mathcal{U}_2 \circ \mathcal{U}_1(\rho_0) + p\rho_1 \\ &= (1-p)^2\mathcal{E}_L \circ \mathcal{U}_L \circ \mathcal{E}_{L-1} \circ \cdots \circ \mathcal{E}_3 \circ \mathcal{U}_3 \circ \mathcal{U}_2 \circ \mathcal{U}_1(\rho_0) + p(1-p)\rho_2 + p\rho_1 \\ &= \cdots \\ &= (1-p)^L\mathcal{U}_{\text{UCC}}(\rho_0) + \sum_{i=1}^L p(1-p)^{i-1}\rho_i, \end{aligned}$$

where, for each $1 \leq i \leq L$, ρ_i is a quantum state defined as

$$\rho_i := (\mathcal{E}_L \circ \mathcal{U}_L) \circ (\mathcal{E}_{L-1} \circ \mathcal{U}_{L-1}) \circ \cdots \circ (\mathcal{E}_{i+1} \circ \mathcal{U}_{i+1}) \circ \Delta_i \circ \mathcal{U}_i \circ \mathcal{U}_{i-1} \circ \cdots \circ \mathcal{U}_2 \circ \mathcal{U}_1(\rho_0)$$

with the completely depolarizing channel Δ_i on the system on which $\mathcal{E}_i$ acts. Then, we have the following chain of inequalities:

$$\begin{aligned} \epsilon_{\text{tot}} &= \|\mathcal{N}_{\text{UCC}}(\rho_0) - \mathcal{U}_{\text{UCC}}(\rho_0)\|_1 \\ &= \left\| \left[(1-p)^L - 1 \right] \mathcal{U}_{\text{UCC}}(\rho_0) + \sum_{i=1}^L p(1-p)^{i-1}\rho_i \right\|_1 \\ &\leq 1 - (1-p)^L + \sum_{i=1}^L p(1-p)^{i-1} \\ &= 2 \left(1 - (1-p)^L \right) \\ &\sim L \times O(\epsilon^2). \end{aligned} \tag{12}$$

Therefore, since $p = O(\epsilon^2)$ for the averaged Clifford+ T approximation error where ϵ is Clifford+ T approximation error for a single-qubit z -rotation [Eq. (18)], the total error ϵ_{tot} caused by Clifford+ T approximation error of each gate can be modeled by a function

$$2 \left(1 - \left(1 - O(\epsilon^2) \right)^L \right).$$

Since the total Clifford+ T approximation error ϵ_{tot} above is given in the trace norm, it is also easy to estimate the error of an expectation value for some observable. Suppose that we evaluate the energy of a given Hamiltonian H with the UCC ansatz. When the ansatz circuit is “noisy”, or approximated, due to the Clifford+ T decomposition, the energy $\tilde{E}$ we obtain under such situation is given by

$$E_{\text{approx}} := \text{Tr}(\mathcal{N}_{\text{UCC}}(\rho_0)H).$$

The difference of the expectation values of H between the ideal ansatz state,

$$E_{\text{ideal}} = \text{Tr}(\mathcal{U}_{\text{UCC}}(\rho_0)H),$$

and the approximated (Clifford+ T decomposed) state is bounded by

$$\begin{aligned} |E_{\text{approx}} - E_{\text{ideal}}| &= |\text{Tr}(\mathcal{N}_{\text{UCC}}(\rho_0)H) - \text{Tr}(\mathcal{U}_{\text{UCC}}(\rho_0)H)| \\ &\leq \|\mathcal{N}_{\text{UCC}}(\rho_0) - \mathcal{U}_{\text{UCC}}(\rho_0)\|_1 \|H\| \\ &\leq 2\|H\| \left(1 - \left(1 - O(\epsilon^2) \right)^L \right). \end{aligned} \tag{13}$$

The Hölder’s inequality [47] is used at the first inequality. The second inequality follows from Eq. (12). Hence, the error of the energy obtained by the Clifford+ T decomposition is expected to depend also on ϵ^2 as shown in Eq. (13).

Note that the analysis conducted in this subsection is applicable not only to the UCC ansatz, but to a broader class of parametrized quantum circuits, in which only z -rotation gates are parametrized. Since Clifford+ z -rotation gates form a universal gate set, we can convert any given parametrized circuit into this form. Thus, in theory, we may use this analysis for an arbitrary ansatz circuit. Practically, this method can be helpful when the Clifford+ z -rotation expression of a given ansatz circuit can be easily obtained.

Remark 1 (Naive upper bound by triangle inequality). If we do not use the averaging method, we only have a looser upper bound using the triangle inequality. Here, we show the derivation of the naive bound.

Let us consider a UCC ansatz circuit $U_{\text{UCC}} = U_L U_{L-1} \cdots U_1$, where each U_i consists of Clifford gates and only one z -rotation gate. Without the averaging method, the resulting circuit by Clifford + T approximation can be written as

$$\tilde{U}_{\text{UCC}} = \tilde{U}_L \circ \tilde{U}_{L-1} \circ \cdots \circ \tilde{U}_1,$$

where for each $i = 1, 2, \dots, L$, $\tilde{U}_i$ only consists of Clifford gates and satisfies

$$\|\tilde{U}_i - U_i\| \leq \epsilon.$$

In this case, by using the triangle inequality, the total approximation error $\tilde{\epsilon}_{\text{tot}}$ is evaluated

as

$$\begin{aligned}
\tilde{\epsilon}_{\text{tot}} &= \left\| \tilde{U}_{\text{UCC}} \rho_0 \tilde{U}_{\text{UCC}}^\dagger - U_{\text{UCC}} \rho_0 U_{\text{UCC}}^\dagger \right\|_1 \\
&= \left\| \tilde{U}_L \tilde{U}_{L-1} \cdots \tilde{U}_1 \rho_0 \tilde{U}_1^\dagger \cdots \tilde{U}_{L-1}^\dagger \tilde{U}_L^\dagger - U_L U_{L-1} \cdots U_1 \rho_0 U_1^\dagger \cdots U_{L-1}^\dagger U_L^\dagger \right\|_1 \\
&= \left\| \tilde{U}_L \tilde{U}_{L-1} \cdots \tilde{U}_1 \rho_0 \tilde{U}_1^\dagger \cdots \tilde{U}_{L-1}^\dagger \tilde{U}_L^\dagger - U_L \tilde{U}_{L-1} \cdots \tilde{U}_1 \rho_0 \tilde{U}_1^\dagger \cdots \tilde{U}_{L-1}^\dagger U_L^\dagger \right. \\
&\quad + U_L \tilde{U}_{L-1} \cdots \tilde{U}_1 \rho_0 \tilde{U}_1^\dagger \cdots \tilde{U}_{L-1}^\dagger U_L^\dagger - U_L U_{L-1} \tilde{U}_{L-2} \cdots \tilde{U}_1 \rho_0 \tilde{U}_1^\dagger \tilde{U}_{L-2} \cdots U_{L-1}^\dagger U_L^\dagger \\
&\quad + \cdots \\
&\quad \left. + U_L U_{L-1} \cdots \tilde{U}_1 \rho_0 \tilde{U}_1^\dagger \cdots U_{L-1}^\dagger U_L^\dagger - U_L U_{L-1} \cdots U_1 \rho_0 U_1^\dagger \cdots U_{L-1}^\dagger U_L^\dagger \right\|_1 \\
&\leq \left\| \tilde{U}_L \tilde{\rho}_{L-1} \tilde{U}_L^\dagger - U_L \tilde{\rho}_{L-1} U_L^\dagger \right\| \\
&\quad + \sum_{i=0}^{L-2} \left\| (U_L \cdots U_{i+2}) \left(\tilde{U}_{i+1} \tilde{\rho}_i \tilde{U}_{i+1}^\dagger - U_{i+1} \tilde{\rho}_i U_{i+1}^\dagger \right) (U_L \cdots U_{i+2})^\dagger \right\|_1,
\end{aligned}$$

where

$$\tilde{\rho}_i := \begin{cases} \rho_0 & i = 0 \\ \tilde{U}_i \cdots \tilde{U}_1 \rho_0 \tilde{U}_1^\dagger \cdots \tilde{U}_i^\dagger & i = 1, 2, \dots, L-1. \end{cases}$$

Due to the unitary-invariance of the trace norm [47], the last line can be rewritten as

$$\sum_{i=0}^{L-1} \left\| \tilde{U}_{i+1} \tilde{\rho}_i \tilde{U}_{i+1}^\dagger - U_{i+1} \tilde{\rho}_i U_{i+1}^\dagger \right\|_1. \quad (14)$$

Each term in Eq. (14) can be further analyzed as

$$\begin{aligned}
&\left\| \tilde{U}_{i+1} \tilde{\rho}_i \tilde{U}_{i+1}^\dagger - U_{i+1} \tilde{\rho}_i U_{i+1}^\dagger \right\|_1 \\
&= \left\| \tilde{U}_{i+1} \tilde{\rho}_i \tilde{U}_{i+1}^\dagger - U_{i+1} \tilde{\rho}_i \tilde{U}_{i+1}^\dagger + U_{i+1} \tilde{\rho}_i \tilde{U}_{i+1}^\dagger - U_{i+1} \tilde{\rho}_i U_{i+1}^\dagger \right\|_1 \\
&\leq \left\| \tilde{U}_{i+1} \tilde{\rho}_i \tilde{U}_{i+1}^\dagger - U_{i+1} \tilde{\rho}_i \tilde{U}_{i+1}^\dagger \right\|_1 + \left\| U_{i+1} \tilde{\rho}_i \tilde{U}_{i+1}^\dagger - U_{i+1} \tilde{\rho}_i U_{i+1}^\dagger \right\|_1 \\
&\leq \|\tilde{U}_{i+1} - U_{i+1}\| \|\tilde{\rho}_i\|_1 \|\tilde{U}_{i+1}^\dagger\| + \|U_{i+1}\| \|\tilde{\rho}_i\|_1 \|\tilde{U}_{i+1}^\dagger - U_{i+1}^\dagger\| \\
&= \|\tilde{U}_{i+1} - U_{i+1}\| + \|\tilde{U}_{i+1}^\dagger - U_{i+1}^\dagger\| \\
&= 2\|\tilde{U}_{i+1} - U_{i+1}\| \\
&\leq 2\epsilon.
\end{aligned}$$

In the first inequality, we used the triangle inequality, and in the second inequality, we use the fact

$$\|ABC\|_1 \leq \|A\| \|B\|_1 \|C\|$$

for linear operators A, B, C [47]. In the second inequality, we used the fact that

$$\|U\| = 1$$

for any unitary operator U . The last equality follows from

$$\|A\| = \|A^\dagger\|$$

for linear operator A [47]. In summary, with this approach, we have

$$\tilde{\epsilon}_{\text{tot}} \leq \sum_{i=1}^L 2\epsilon = 2L\epsilon. \quad (15)$$

While the averaging method lead to an upper bound with the scaling of $\sim L \times O(\epsilon^2)$, this naive method yields the linear upper bound $2L\epsilon$.

Now, we evaluate the expectation value of H under this approximated ansatz state,

$$\tilde{E}_{\text{approx}} = \text{Tr} \left(\tilde{U}_{\text{UCC}} \rho_0 \tilde{U}_{\text{UCC}}^\dagger H \right). \quad (16)$$

By using the bound (15), in a similar matter to Eq. (13), we have

$$\begin{aligned} |\tilde{E}_{\text{approx}} - E_{\text{ideal}}| &= \left| \text{Tr} \left(\tilde{U}_{\text{UCC}} \rho_0 \tilde{U}_{\text{UCC}}^\dagger H \right) - \text{Tr} \left(U_{\text{UCC}} \rho_0 U_{\text{UCC}}^\dagger H \right) \right| \\ &\leq \left\| \tilde{U}_{\text{UCC}} \rho_0 \tilde{U}_{\text{UCC}}^\dagger - U_{\text{UCC}} \rho_0 U_{\text{UCC}}^\dagger \right\|_1 \|H\| \\ &\leq 2\|H\|L\epsilon. \end{aligned} \quad (17)$$

5.2 Algorithm of T -Count Fixed Gate Decomposition

In this section, we present our algorithm to obtain a decomposition of a given single-qubit z -rotation *with a fixed number of T gates*. In the successive section (Sec. 5.3), we perform numerical simulations by using this algorithm to figure out the performance of the UCC ansatz in the Pre-FTQC era. We expect that our algorithm is useful when one wants to obtain a good Clifford+ T decomposition when the available number of T gates is pre-determined, which should be highly plausible in Pre-FTQC. Our algorithm is based on the one proposed in Ref. [34], which yields a decomposition of a given z -rotation with the *minimal number of T gates* within a *fixed accuracy*.

First, we introduce notations and background materials for the algorithm. Let ω denote a complex number $e^{i\frac{\pi}{4}}$. We define the following rings, which are given as extensions of the ring of the integers $\mathbb{Z}$.

Definition 1. The ring $\mathbb{Z}[\omega]$ is defined by

$$\mathbb{Z}[\omega] := \left\{ a\omega^3 + b\omega^2 + c\omega + d \mid a, b, c, d \in \mathbb{Z} \right\}.$$

The ring $\mathbb{D}$ is defined by

$$\mathbb{D} := \left\{ \frac{a}{2^k} \mid a \in \mathbb{Z}, k = 0, 1, 2, \dots \right\}.$$

The ring $\mathbb{D}[\omega]$ is defined by

$$\mathbb{D}[\omega] := \left\{ a\omega^3 + b\omega^2 + c\omega + d \mid a, b, c, d \in \mathbb{D} \right\}.$$

Definition 2. Let $t \in \mathbb{D}[\omega]$ and $k \in \mathbb{N}$. If $\sqrt{2^k}t \in \mathbb{Z}[\omega]$, then we say that k is a *denominator exponent* of t . The smallest such $k \geq 0$ is the *least denominator exponent* of t .

Now, let us describe our gate decomposition algorithm. Due to Lemmas 7.2 and 7.3 of Ref. [34], the task of finding a good approximation of a given z -rotation $R_z(\theta) = e^{-i\frac{\theta}{2}Z}$ within a given number N_T of T -gates can be reduced to the following searching problem.

Problem 1. Let $\theta \in [0, 2\pi)$ be a given angle of z -rotation, and let N_T be a positive integer representing the maximum number of T gates available. Find a pair of $t, u \in \mathbb{D}[\omega]$ maximizing $\text{Re}(ue^{-i\theta/2})$ under the following constraints:

- (a) $tt^\dagger + uu^\dagger = 1$,
- (b) and u has a denominator exponent $\lfloor \frac{N_T}{2} \rfloor + 1$.

Once a pair of (t, u) are obtained, the unitary matrix

$$U(t, u) := \begin{pmatrix} u & -t^\dagger \\ t & u^\dagger \end{pmatrix}$$

serves as a good approximation of the given z -rotation $R_z(\theta)$. There exists an efficient algorithm to obtain the exact decomposition of $U(t, u)$ into H , S , and T gates [48]. In this exact decomposition algorithm, the number of T gates used is given by $2k - 2$ with the least denominator exponent k of u when $k > 0$, and T -count is zero when $k = 0$. Simple calculation shows that Clifford+ T approximation error is indeed given by

$$\epsilon := \|U(t, u) - R_z(\theta)\| = \sqrt{2(1 - \text{Re}(ue^{-i\theta/2}))}, \quad (18)$$

where $\|\cdot\|$ is the operator norm.

Now, we show an algorithm to have the best approximate decomposition of a given z rotation within a given cost of T -gates.

Algorithm 1. Suppose that an angle θ , the maximum number N_T of T gates, and an estimate of Clifford+ T approximation error for each gate $\tilde{\epsilon} > 0$ are given. By using Lemma 7.4 and Proposition 5.21 of Ref. [34], we can efficiently enumerate the candidates of $u \in \mathbb{D}[\omega]$ with the denominator exponent $\lfloor \frac{N_T}{2} \rfloor + 1$ that satisfy $tt^\dagger + uu^\dagger = 1$ and $\text{Re}(ue^{-i\theta/2}) \geq 1 - \tilde{\epsilon}^2/2$. Let $\mathbb{U}_{\text{cand}}$ be the set of such candidates u . If no such u was found, that is, $\mathbb{U}_{\text{cand}} = \emptyset$, we set $\tilde{\epsilon} \leftarrow 2\tilde{\epsilon}$, and try with this new error estimate.

- (i) Set $\epsilon_{\text{tmp}} = 2$, and for each candidate $u \in \mathbb{U}_{\text{cand}}$, we run the following procedure:
Use Theorem 6.2 of Ref. [34] to solve $tt^\dagger = 1 - uu^\dagger$ with respect to t . If there is no solution or $\epsilon = \sqrt{2 - 2\text{Re}(ue^{-i\theta/2})} \geq \epsilon_{\text{tmp}}$, then skip to the next candidate in $\mathbb{U}_{\text{cand}}$. Otherwise, set

$$t_{\text{tmp}} \leftarrow t, \quad u_{\text{tmp}} \leftarrow u, \quad \epsilon_{\text{tmp}} \leftarrow \epsilon \quad (19)$$

and continue with the next u .

- (ii) After examining all the candidates, use the exact synthesis algorithm of Ref. [48] to find decomposition of $U(t_{\text{tmp}}, u_{\text{tmp}})$ into H , S , and T gates.
- (iii) Output the circuit $U(t_{\text{tmp}}, u_{\text{tmp}})$ and Clifford+ T approximation error ϵ_{tmp} , and stop.

Remark 2 (Error Estimate $\tilde{\epsilon}$). Algorithm 1 searches all the Clifford+ T -gate decompositions within estimate error $\tilde{\epsilon}$, and then outputs the one with the smallest error. Thus, in theory, a sufficiently large $\tilde{\epsilon}$, say, $\tilde{\epsilon} = 2$, guarantees the smallest-error decomposition is obtained through this algorithm. Practically, smaller $\tilde{\epsilon}$ may be preferable in terms of (classical) computational efficiency since the search region $\mathbb{U}_{\text{cand}}$ is smaller when we take smaller ϵ . Hence, we consider inputting some small enough error estimate or desired accuracy as an initial $\tilde{\epsilon}$, and updating $\tilde{\epsilon} \leftarrow 2\tilde{\epsilon}$ when no decomposition is obtained. Since the algorithm returns the smallest-error decomposition anyway, the choice of the initial estimated error will not affect the validity of the algorithm.

Remark 3 (Comparison with the algorithm in Ref. [34]). The algorithm proposed in Ref. [34] provides a Clifford+ T -gate decomposition with the minimum T -count under given Clifford+ T approximation error ϵ . Thus, in that algorithm, one enumerates the candidate of u with the smallest possible denominator exponent k , say $k = 0$, that satisfies $tt^\dagger + uu^\dagger = 1$ and $\text{Re}(ue^{-i\theta/2}) \geq 1 - \epsilon^2$ with the fixed given ϵ . If no such pair (t, u) was found, then one increments the denominator exponent k by 1, and repeat the same procedure until such a pair is found. Once a pair of (t, u) is found, then one produces the corresponding Clifford+ T -gate decomposition using the exact synthesis algorithm [48]. The primary difference between these two algorithms is that we fix N_T and seek the best ϵ while they fixed ϵ and sought the best N_T .

Remark 4 (Necessity of Prime Factorization). Similar to the algorithm by Ross and Selinger [34], our algorithm demands prime factorization of some integer when we solve $tt^\dagger + uu^\dagger = 1$. Although factoring is in NP, for which a computationally efficient algorithm has not been found, factoring in this algorithm can be done fast practically by the brute-force method since the integer which we want to perform factorization is relatively small, as also indicated in Ref. [34].

5.3 Numerical experiments for the UCC ansatz

Here we numerically illustrate the theoretical analysis on the average Clifford+ T approximation error above by actually decomposing the quantum circuit of the UCC ansatz using the algorithm described in Sec. 5.2.

Our setup for numerical simulations is as follows. We consider a wide variety of molecules shown in Table 1. We adopt the STO-3G minimal basis set, and construct the fermionic second-quantized Hamiltonians for electrons with the Hartree-Fock orbitals by using the numerical libraries PySCF [49] and OpenFermion [50]. We then employ Jordan-Wigner transformation [51] to map the fermionic Hamiltonians into qubit Hamiltonians [52, 53]. We consider decomposing the unitary coupled cluster singles and doubles (UCCSD) ansatz circuit for those molecules, which is explicitly defined in Appendix B, into Clifford+ T -gate circuit by using the algorithm in Sec. 5.2. In the simulation, we do not optimize the parameters in the UCCSD ansatz. We instead use the amplitudes of the coupled cluster singles and doubles (CCSD) [45] computed on classical computers with PySCF [49] as parameters of the UCCSD ansatz. Hereafter, we call this state the *UCCSD' state*. We generate Clifford+ T -gate decomposition of the UCC' ansatz for each molecule with various T -count N_T , which stands for the maximum number of T -gates used in Algorithm 1 in Sec. 5.2. We consider $22 \leq N_T \leq 50$ in the simulation. All simulations for quantum circuits are performed by the high-speed quantum circuit simulator Qulacs [54]. We do not include any noise for quantum circuit simulations.

First, we examine the Clifford+ T approximation error analysis proposed in Sec. 4. We calculate the energy expectation values of the approximate UCCSD' state under Clifford+ T decomposition with the single-gate T -count N_T . Recalling that the average effects of Clifford+ T approximation error on the UCC ansatz is given by Eq. (13), we model the energy expectation value for a given N_T as

$$E_{\text{ideal}} + 2E_{\text{ideal}} \left(1 - \left(1 - c \cdot (\epsilon(N_T))^2 \right)^L \right), \quad (20)$$

where $\epsilon(N_T) := 10^{-\frac{N_T}{10}}$ is the single-gate Clifford+ T approximation error, L is the total number of z rotations in the UCCSD ansatz, E_* is the energy expectation value for the original UCCSD ansatz (not decomposed into Clifford+ T), and c is a fitting parameter.

Table 1: Geometries of molecules used in the numerical simulations. “(X, (x, y, z))” denotes three dimensional coordinates x, y, z of atom X in units of Å.

Molecule	# of Qubits	Geometry
H ₂	4	(H, (0, 0, 0)), (H, (0, 0, 0.735))
H _{<i>i</i>} ($i = 4, 6, 8, 10$)	2 <i>i</i>	(H, (0, 0, $j - 1$)) for the j th H atom ($j = 1, 2, \dots, i$).
H ₂ O	14	(O, (0, 0, 0.137)), (H, (0, 0.769, -0.546)), (H, (0, -0.769, -0.546))
LiH	12	(Li, (0, 0, 0)), (H, (0, 0, 1.548))
BeH ₂	14	(Be, (0, 0, 0)), (H, (0, 0, -1.304)), (H, (0, 0, 1.304))
H ₂ S	22	(S, (0, 0, 0.1030)), (H, (0, 0.9616, -0.8239)), (H, (0, -0.9616, -0.8239))
CH ₄	18	(C, (0, 0, 0)), (H, (0.6276, 0.6276, 0.6276)), (H, (0.6276, -0.6276, -0.6276)), (H, (-0.6276, 0.6276, -0.6276)), (H, (-0.6276, -0.6276, 0.6276))
C ₂ H ₂	24	(C, (0, 0, 0.6013)), (C, (0, 0, -0.6013)), (H, (0, 0, 1.6644)), (H, (0, 0, -1.6644))
CO	20	(C, (0, 0, 0)), (O, (0, 0, 1.128))
N ₂	20	(N, (0, 0, 0)), (N, (0, 0, 1.5))
NH ₃	16	(N, (0, 0, 0.149)), (H, (0, 0.947, -0.348)), (H, (0.821, -0.474, -0.348)), (H, (-0.821, -0.474, -0.348))

The functional form of the single-gate Clifford+ T approximation error $10^{-\frac{N_T}{10}}$ is assumed by following the numerical observation in Ref. [34]. For each molecule, the data of the energy expectation values versus N_T is fitted by performing the (non-linear) regression to this model (20). We take the range of $32 \leq N_T \leq 50$ for the fitting to see the accuracy of the model for relatively large N_T . The results for C₂H₂ and NH₃ molecules, which exhibit representative behaviors, are shown in Fig. 1. The whole results for all molecules in Table 1 are given in Appendix C. As one may see, the model (20) explains the total Clifford+ T approximation error well while the previous error estimate by the triangle inequality gives much worse upper bound of the energy (see the captions). This is an illustration of the validity of the average error analysis presented in Sec. 4. We also compare our results with the naive estimation using the triangle inequality (17) (see the caption and subcaptions of Fig. 1). Note that while $\|H\|$ cannot be obtained when a given quantum system is large, we here use the true value of the ground energy. The comparison indeed shows that our model gives a more precise estimation of energy expectation values.

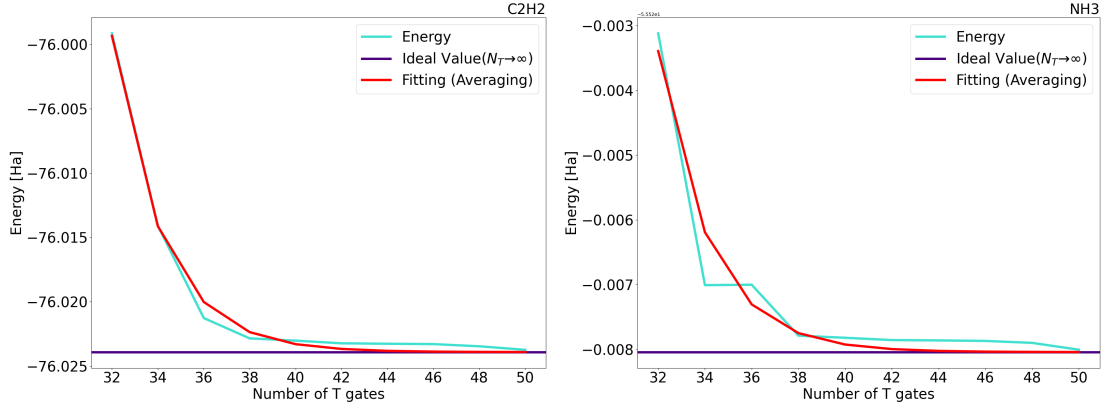
Next, we evaluate the threshold for the T -count N_T with which we can guarantee sufficient accuracy to represent the UCCSD’ state. In Fig. 2, we plot the fidelity between the exact ground state of the Hamiltonian and the approximate (Clifford+ T -gate-decomposed) UCCSD’ state with respect to the T -count. As expected, the larger N_T results in higher fidelity. We determine the threshold for T -count to represent the UCCSD’ state with high accuracy by requiring that the following quantity,

$$\delta_{\text{rel}} := \frac{|F(N_T) - F_{\text{ideal}}|}{F_{\text{ideal}}}, \quad (21)$$

satisfies $\delta_{\text{rel}} < 10^{-4}$, where $F(N_T)$ is the fidelity between the approximated UCCSD’ state and the exact ground state of the Hamiltonian and F_{ideal} is the fidelity between the original (not-decomposed) UCCSD’ state and the exact ground state. δ_{del} means the relative error of the fidelity with respect to the exact ground state of the Hamiltonian. The results for

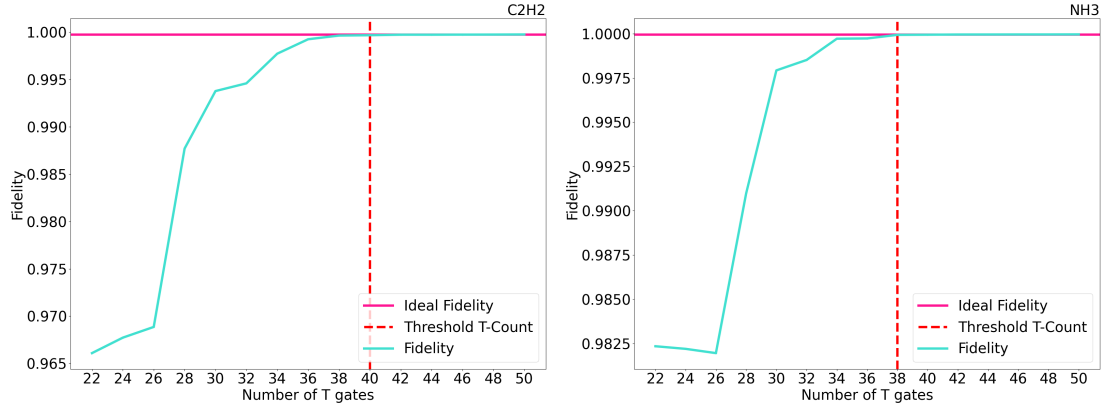
the *total* threshold T -counts, ((the per-gate threshold T -counts obtained in Fig. 2) $\times$ (the number of z -rotation gates)), for various molecules are summarized in Fig. 3. Indeed, the total threshold T -count mostly increases quartically with respect to the number of qubits. In Fig. 3, we also show the results of fitting of the total threshold T -count versus the number of qubits obtained by performing the linear regression using a quartic function an^4 with fitting parameter a and the number n of qubits. The fitting explains the data pretty well with fitting result $a = 0.07795(1)$. This n^4 dependence is natural when one considers there are $O(n^4)$ multi-qubit Pauli rotation gates in the UCCSD ansatz.

This result implies that the usage of variational quantum algorithms with a deep ansatz such as the UCC ansatz can be beneficial in the Pre-FTQC era. For example, consider the FeMo cofactor of the nitrogenase enzyme $\text{Fe}_7\text{MoS}_9\text{C}$, known as FeMoco. There are two active space models considered conventionally for this molecule, one proposed by Reiher *et al.* [38] with the number of qubit $n = 104$ (RWSWT orbital) and the other one proposed by Li *et al.* [55] with $n = 154$ (LLDUC orbital). The current best resource estimation for the quantum phase estimation of FeMoco, as far as we are aware of, is obtained using tensor hypercontraction by Lee *et al.* [42], and the estimate *Toffoli* count is 5.3×10^9 for RWSWT FeMoco and 3.2×10^{10} for LLDUC FeMoco. The most straightforward way to estimate T -count from *Toffoli* count is to simply transform the *Toffoli* gate into a quantum circuit using Clifford+ T -gates, and this can be done using 7 T gates [13]. Or, recently, a magic-state factory construction based on the surface code was proposed [56], implying one *Toffoli* gate is roughly twice as costly as one T gate. With these estimations, for both active space models, the T -count needed to perform the quantum phase estimation to FeMoco is roughly $O(10^{10}) \sim O(10^{11})$ at best. On the other hand, with our estimation, the T -count for realizing the UCCSD' state with the accuracy $\delta_{\text{rel}} < 10^{-4}$ is $an^4 \approx 0.078 \times (154)^4 \approx 4 \times 10^7$ even for the larger active space of LLDUC FeMoco, which is smaller than the estimate T -count for the quantum phase estimation by three orders of magnitude. In addition, the tensor hypercontraction demands enormous auxiliary qubits [42]; if we do not use auxiliary qubits, adopting the Trotter-based methods, we have the estimate T -count of roughly $O(10^{14}) \sim O(10^{16})$. In this case, the T -count for the UCCSD' state is even smaller by seven-to-nine orders of magnitude. During the Pre-FTQC era, it is expected that we will experience time where we have $O(10^7)$ T gates but still cannot perform the phase estimation algorithm due to the lack of a sufficient number of clean qubits, an possibly improved but still high error rate, and the limited number of T gates. Our numerical results imply the Clifford+ T approximated UCCSD ansatz may be used to generate a good approximation of a desired state in such an era. Moreover, even in the *post* Pre-FTQC era, where we are stepping into FTQC and the phase estimation algorithm is finally available, the Clifford+ T approximated UCCSD ansatz may be useful to generate a good initial state for the phase estimation considering the T -counts. Thus, the averaging method we proposed and our numerical analysis strongly suggest the utility of the combination of the Clifford+ T approximation and parametrized ansatz states in the Pre-FTQC era and beyond.



(a) The energy expectation values of C₂H₂. The fitting parameter is $c = 0.6106(2)$. On the other hand, the triangle inequality gives energy upper bound -75.52 Ha at $N_T = 50$. (b) The energy expectation values of NH₃. The fitting parameter is $c = 0.779(2)$. The triangle inequality gives energy upper bound -55.45 Ha at $N_T = 50$.

Figure 1: Energy expectation values of Clifford+ T -gate decomposed UCCSD' states with T -gate budget $32 \leq N_T \leq 50$. In each panel, the orange curve represents the expectation values of the UCCSD' state that are decomposed by Algorithm 1. The purple horizontal line indicates the true UCCSD' state energy obtained by the UCCSD ansatz without Clifford+ T decomposition. The blue curve is the fitting of the orange curve modeled by Eq. (20). As indicated in each subcaption, the naive analysis by the triangle inequality is substantially rougher than our method.



(a) Fidelity of the UCCSD' states of C₂H₂ to the exact ground state. (b) Fidelity of the UCCSD' states of NH₃ to the exact ground state.

Figure 2: The fidelity between the exact ground state of the Hamiltonian and the Clifford+ T -gate decomposed UCCSD' states with T -gate budget $22 \leq N_T \leq 50$. In each panel, the blue curve represents the fidelity between the exact ground state and the UCCSD' state that is decomposed by Algorithm 1. The pink horizontal line indicates the fidelity between the exact ground state and the true UCCSD' state obtained by the UCCSD ansatz without Clifford+ T decomposition. The orange vertical line represents the T -count threshold that achieves Eq. (21).

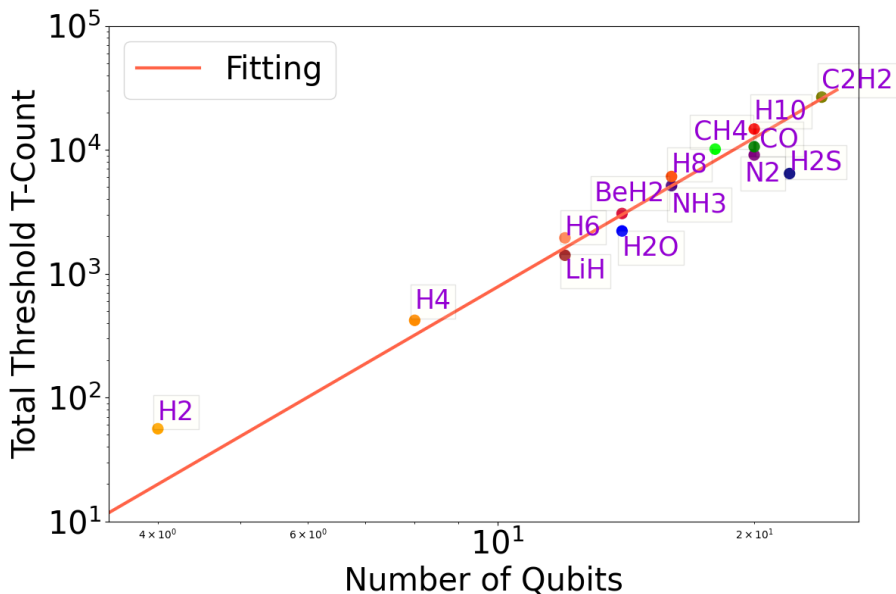


Figure 3: Number of qubits vs *Total* threshold *T*-count. The red curve is the fitting of the data of the total threshold *T*-count versus the number of qubits by a quartic function an^4 with the fitting result $a = 0.07795(1)$, where n is the number of qubits.

6 Summary discussion, and outlook

In this work, we proposed a method of averaging the effects of the unitary Clifford+*T* approximation error into the depolarizing noise when a given circuit is sufficiently large, which makes the analysis of the total Clifford+*T* approximation error easier. We apply this method to Unitary Coupled Cluster (UCC) ansatz. We constructed a Clifford+*T*-gate decomposition algorithm to provide the most accurate approximation of a given single-qubit *z*-rotation gate when the available number of *T* gates is restricted. With this algorithm, we numerically verified our proposal for a wide variety of molecules, and based on the numerical simulations, we reveal the threshold *T*-counts to achieve desired accuracy. Hence, we shed a new light for the Pre-FTQC era, where only a restricted number of *T* gates are practically available. We believe that our proposal and analysis lead to more accurate resource estimation in the framework of Pre-FTQC. As a future direction, it would be nice if our algorithm and averaging method are applicable to another setup, *e.g.*, quantum phase estimation algorithm, and estimate Clifford+*T* approximation error or threshold *T*-count in that case. Moreover, it would also be interesting to investigate the performance of a Clifford+*T* decomposed ansatz in a variational quantum algorithm, where the number of available *T* gates is limited. In our numerical experiment, we compared the energy expectation values of the UCCSD' states with the energies computed by the classical CCSD method to verify the performance of our theoretical analysis. However, one may want to try to obtain a good estimate of the ground energy of a given molecule by a variational algorithm, in particular, the variational quantum eigensolver (VQE) with a Clifford+*T* approximated ansatz with a limited number of *T* gates in the Pre-FTQC era. By applying our analysis to the ground state, we believe that the resulting energy estimate obtained by the VQE with a Clifford+*T* approximated ansatz behaves similarly to what we obtained in our numerical simulation. It would be great if this observation is verified in future research.

Note added

After posting the previous version of this manuscript on arXiv, during the preparation of this version of the manuscript, we noticed the appearance of a few papers relevant to our results. In Ref. [57], Akibue *et al.* proved that gate approximation error becomes in the order of $O(\epsilon^2)$ when one performs gate approximation probabilistically, that is, randomly selecting an approximated gate from an appropriate set of candidates with appropriate probability distribution. In Ref. [58], Yoshioka *et al.* proved that one can craft the remnant error of the probabilistic approximation of the target unitary into a Pauli channel, which is useful to apply the error countermeasure such as quantum error mitigation, while keeping the error $O(\epsilon^2)$. They also numerically observed that the approximation error might drop $O(\epsilon^3)$ when one allows CPTP maps with ancillary qubits.

In this paper, contrary to their results, we do not consider probabilistic approximation or any countermeasure against error. Rather, we assume that when a given parametrized circuit is sufficiently large, the effect of gate approximation error for each gate will be averaged over the whole circuit. Under this setup, we observe that the total error can be model as the depolarizing channel with error $O(\epsilon^2)$. It would be interesting to see if there is a deeper connection between our analysis and results obtained by Refs. [57, 58].

Acknowledgement

The authors are grateful to Yasunori Lee for valuable discussions and comments on the manuscript. We also thank Keita Kanno for fruitful discussions on the early stage of this work. We would like to express our deep appreciation to Hayata Morisaki for providing us with useful codes of the Clifford+ T decomposition algorithm of Ref. [34]. The numerical simulations of this work are largely inspired by his implementation.

KK is supported by a Mike and Ophelia Lazaridis Fellowship, the Funai Foundation, and a Perimeter Residency Doctoral Award.

References

- [1] Peter W. Shor. “Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer”. *SIAM J. Comput.* **26**, 1484–1509 (1997).
- [2] Lov K. Grover. “A Fast Quantum Mechanical Algorithm for Database Search”. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*. Page 212–219. STOC '96 New York, NY, USA (1996). Association for Computing Machinery.
- [3] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen

- Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Viallonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. “Quantum supremacy using a programmable superconducting processor”. *Nature* **574**, 505–510 (2019).
- [4] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu, Peng Hu, Xiao-Yan Yang, Wei-Jun Zhang, Hao Li, Yuxuan Li, Xiao Jiang, Lin Gan, Guangwen Yang, Lixing You, Zhen Wang, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan. “Quantum computational advantage using photons”. *Science* **370**, 1460–1463 (2020).
- [5] Han-Sen Zhong, Yu-Hao Deng, Jian Qin, Hui Wang, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Dian Wu, Si-Qiu Gong, Hao Su, Yi Hu, Peng Hu, Xiao-Yan Yang, Wei-Jun Zhang, Hao Li, Yuxuan Li, Xiao Jiang, Lin Gan, Guangwen Yang, Lixing You, Zhen Wang, Li Li, Nai-Le Liu, Jelmer J. Renema, Chao-Yang Lu, and Jian-Wei Pan. “Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light”. *Phys. Rev. Lett.* **127**, 180502 (2021).
- [6] Yulin Wu, Wan-Su Bao, Sirui Cao, Fusheng Chen, Ming-Cheng Chen, Xiawei Chen, Tung-Hsun Chung, Hui Deng, Yajie Du, Daojin Fan, Ming Gong, Cheng Guo, Chu Guo, Shaojun Guo, Lianchen Han, Linyin Hong, He-Liang Huang, Yong-Heng Huo, Liping Li, Na Li, Shaowei Li, Yuan Li, Futian Liang, Chun Lin, Jin Lin, Haoran Qian, Dan Qiao, Hao Rong, Hong Su, Lihua Sun, Liangyuan Wang, Shiyu Wang, Dachao Wu, Yu Xu, Kai Yan, Weifeng Yang, Yang Yang, Yangsen Ye, Jianghan Yin, Chong Ying, Jiale Yu, Chen Zha, Cha Zhang, Haibin Zhang, Kaili Zhang, Yiming Zhang, Han Zhao, Youwei Zhao, Liang Zhou, Qingling Zhu, Chao-Yang Lu, Cheng-Zhi Peng, Xiaobo Zhu, and Jian-Wei Pan. “Strong Quantum Computational Advantage Using a Superconducting Quantum Processor”. *Phys. Rev. Lett.* **127**, 180501 (2021).
- [7] Andrew J. Daley, Immanuel Bloch, Christian Kokail, Stuart Flannigan, Natalie Pearson, Matthias Troyer, and Peter Zoller. “Practical quantum advantage in quantum simulation”. *Nature* **607**, 667–676 (2022).
- [8] Lars S. Madsen, Fabian Laudenbach, Mohsen Falamarzi, Askarani, Fabien Rortais, Trevor Vincent, Jacob F. F. Bulmer, Filippo M. Miatto, Leonhard Neuhaus, Lukas G. Helt, Matthew J. Collins, Adriana E. Lita, Thomas Gerrits, Sae Woo Nam, Varun D. Vaidya, Matteo Menotti, Ish Dhand, Zachary Vernon, Nicolás Quesada, and Jonathan Lavoie. “Quantum computational advantage with a programmable photonic processor”. *Nature* **606**, 75–81 (2022).
- [9] John Preskill. “Quantum Computing in the NISQ era and beyond”. *Quantum* **2**, 79 (2018).
- [10] A. Yu. Kitaev. “Quantum measurements and the abelian stabilizer problem” (1995). [arXiv:quant-ph/9511026](https://arxiv.org/abs/quant-ph/9511026).
- [11] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca. “Quantum algorithms revisited”. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **454**, 339–354 (1998).
- [12] Daniel Gottesman. “An introduction to quantum error correction and fault-tolerant quantum computation.”. In *Quantum Information Science and Its Contributions to Mathematics, Proceedings of Symposia in Applied Mathematics*. Volume 68, pages 13–58. American Mathematical Society (2010).
- [13] Michael A. Nielsen and Isaac L. Chuang. “Quantum Computation and Quantum Information: 10th Anniversary Edition”. *Cambridge University Press*. (2010).

- [14] A Yu Kitaev. “Quantum computations: algorithms and error correction”. *Russian Math. Surveys* **52**, 1191 (1997).
- [15] Christopher M. Dawson and Michael A. Nielsen. “The Solovay-Kitaev algorithm” (2005). [arXiv:quant-ph/0505030](#).
- [16] P.W. Shor. “Fault-tolerant quantum computation”. In Proceedings of 37th Conference on Foundations of Computer Science. Pages 56–65. (1996).
- [17] A. M. Steane. “Active Stabilization, Quantum Computation, and Quantum State Synthesis”. *Phys. Rev. Lett.* **78**, 2252–2255 (1997).
- [18] E. Knill. “Quantum computing with realistically noisy devices”. *Nature* **434**, 39–44 (2005).
- [19] Daniel Gottesman. “Theory of fault-tolerant quantum computation”. *Phys. Rev. A* **57**, 127–137 (1998).
- [20] Daniel Gottesman and Isaac L. Chuang. “Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations”. *Nature* **402**, 390–393 (1999).
- [21] Xinlan Zhou, Debbie W. Leung, and Isaac L. Chuang. “Methodology for quantum logic gate construction”. *Phys. Rev. A* **62**, 052316 (2000).
- [22] Sergey Bravyi and Alexei Kitaev. “Universal quantum computation with ideal clifford gates and noisy ancillas”. *Phys. Rev. A* **71**, 022316 (2005).
- [23] D Gottesman. “The Heisenberg representation of quantum computers”. In S. P. Corney, R. Delbourgo, and P. D. Jarvis, editors, Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics. Pages 32–43. (1998). [arXiv:quant-ph/9807006](#).
- [24] Scott Aaronson and Daniel Gottesman. “Improved simulation of stabilizer circuits”. *Phys. Rev. A* **70**, 052328 (2004).
- [25] Lin Lin and Yu Tong. “Heisenberg-Limited Ground-State Energy Estimation for Early Fault-Tolerant Quantum Computers”. *PRX Quantum* **3**, 010318 (2022).
- [26] Rutuja Kshirsagar, Amara Katabarwa, and Peter D. Johnson. “On proving the robustness of algorithms for early fault-tolerant quantum computers”. *Quantum* **8**, 1531 (2024).
- [27] Yasunari Suzuki, Suguru Endo, Keisuke Fujii, and Yuuki Tokunaga. “Quantum Error Mitigation as a Universal Error Reduction Technique: Applications from the NISQ to the Fault-Tolerant Quantum Computing Eras”. *PRX Quantum* **3**, 010345 (2022).
- [28] Zhiyan Ding and Lin Lin. “Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation”. *PRX Quantum* **4**, 020331 (2023).
- [29] Hayata Morisaki, Kosuke Mitarai, Keisuke Fujii, and Yuya O. Nakagawa. “Classical variational optimization of a prepare circuit for quantum phase estimation of quantum chemistry hamiltonians”. *Phys. Rev. Res.* **6**, 043186 (2024).
- [30] Hasan Sayginel, Francois Jamet, Abhishek Agarwal, Dan E Browne, and Ivan Rungger. “A fault-tolerant variational quantum algorithm with limited t-depth”. *Quantum Science and Technology* **9**, 015015 (2023).
- [31] Amara Katabarwa, Katerina Gratsea, Athena Caesura, and Peter D. Johnson. “Early fault-tolerant quantum computing”. *PRX Quantum* **5**, 020101 (2024).
- [32] Yutaro Akahoshi, Kazunori Maruyama, Hirotaka Oshima, Shintaro Sato, and Keisuke Fujii. “Partially fault-tolerant quantum computing architecture with error-corrected clifford gates and space-time efficient analog rotations”. *PRX Quantum* **5**, 010337 (2024).

- [33] Abhinav Anand, Philipp Schleich, Sumner Alperin-Lea, Phillip W. K. Jensen, Sukin Sim, Manuel Díaz-Tinoco, Jakob S. Kottmann, Matthias Degroote, Artur F. Izmaylov, and Alán Aspuru-Guzik. “A quantum computing view on unitary coupled cluster theory”. *Chem. Soc. Rev.* **51**, 1659–1684 (2022).
- [34] Neil J. Ross and Peter Selinger. “Optimal Ancilla-Free Clifford+T Approximation of z-Rotations”. *Quantum Info. Comput.* **16**, 901–953 (2016). [arXiv:1403.2975](#).
- [35] Earl Campbell. “Shorter gate sequences for quantum computing by mixing unitaries”. *Phys. Rev. A* **95**, 042306 (2017).
- [36] Zhenyu Cai, Ryan Babbush, Simon C. Benjamin, Suguru Endo, William J. Huggins, Ying Li, Jarrod R. McClean, and Thomas E. O’Brien. “Quantum error mitigation”. *Rev. Mod. Phys.* **95**, 045005 (2023).
- [37] Lorenza Viola, Emanuel Knill, and Seth Lloyd. “Dynamical decoupling of open quantum systems”. *Phys. Rev. Lett.* **82**, 2417–2421 (1999).
- [38] Markus Reiher, Nathan Wiebe, Krysta M. Svore, Dave Wecker, and Matthias Troyer. “Elucidating reaction mechanisms on quantum computers”. *Proceedings of the National Academy of Sciences* **114**, 7555–7560 (2017).
- [39] Ryan Babbush, Craig Gidney, Dominic W. Berry, Nathan Wiebe, Jarrod McClean, Alexandru Paler, Austin Fowler, and Hartmut Neven. “Encoding Electronic Spectra in Quantum Circuits with Linear T Complexity”. *Phys. Rev. X* **8**, 041015 (2018).
- [40] Dominic W. Berry, Craig Gidney, Mario Motta, Jarrod R. McClean, and Ryan Babbush. “Qubitization of Arbitrary Basis Quantum Chemistry Leveraging Sparsity and Low Rank Factorization”. *Quantum* **3**, 208 (2019).
- [41] Vera von Burg, Guang Hao Low, Thomas Häner, Damian S. Steiger, Markus Reiher, Martin Roetteler, and Matthias Troyer. “Quantum computing enhanced computational catalysis”. *Phys. Rev. Research* **3**, 033055 (2021).
- [42] Joonho Lee, Dominic W. Berry, Craig Gidney, William J. Huggins, Jarrod R. McClean, Nathan Wiebe, and Ryan Babbush. “Even More Efficient Quantum Computations of Chemistry Through Tensor Hypercontraction”. *PRX Quantum* **2**, 030305 (2021).
- [43] Nobuyuki Yoshioka, Tsuyoshi Okubo, Yasunari Suzuki, Yuki Koizumi, and Wataru Mizukami. “Hunting for quantum-classical crossover in condensed matter problems”. *npj Quantum Information* **10**, 45 (2024).
- [44] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O’Brien. “A variational eigenvalue solver on a photonic quantum processor”. *Nature Communications* **5**, 4213 (2014).
- [45] Rodney J. Bartlett and Monika Musiał. “Coupled-cluster theory in quantum chemistry”. *Rev. Mod. Phys.* **79**, 291–352 (2007).
- [46] Seunghoon Lee, Joonho Lee, Huanchen Zhai, Yu Tong, Alexander M. Dalzell, Ashutosh Kumar, Phillip Helms, Johnnie Gray, Zhi-Hao Cui, Wenyan Liu, Michael Kastoryano, Ryan Babbush, John Preskill, David R. Reichman, Earl T. Campbell, Edward F. Valeev, Lin Lin, and Garnet Kin-Lic Chan. “Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry”. *Nature Communications* **14**, 1952 (2023).
- [47] John Watrous. “The Theory of Quantum Information”. *Cambridge University Press*. (2018).
- [48] Vadym Kliuchnikov, Dmitri Maslov, and Michele Mosca. “Fast and Efficient Exact Synthesis of Single-Qubit Unitaries Generated by Clifford and T Gates”. *Quantum Info. Comput.* **13**, 607–630 (2013). [arXiv:1206.5236](#).

- [49] Qiming Sun, Timothy C. Berkelbach, Nick S. Blunt, George H. Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D. McClain, Elvira R. Sayfutyarova, Sandeep Sharma, Sebastian Wouters, and Garnet Kin-Lic Chan. “PySCF: the Python-based simulations of chemistry framework”. *WIREs Comput. Mol. Sci.* **8**, e1340 (2018).
- [50] Jarrod R McClean, Nicholas C Rubin, Kevin J Sung, Ian D Kivlichan, Xavier Bonet-Monroig, Yudong Cao, Chengyu Dai, E Schuyler Fried, Craig Gidney, Brendan Gimby, Pranav Gokhale, Thomas Häner, Tarini Hardikar, Vojtěch Havlíček, Oscar Higgott, Cupjin Huang, Josh Izaac, Zhang Jiang, Xinle Liu, Sam McArdle, Matthew Neeley, Thomas O’Brien, Bryan O’Gorman, Isil Ozfidan, Maxwell D Radin, Jhonathan Romero, Nicolas P D Sawaya, Bruno Senjean, Kanav Setia, Sukin Sim, Damian S Steiger, Mark Steudtner, Qiming Sun, Wei Sun, Daochen Wang, Fang Zhang, and Ryan Babbush. “Openfermion: the electronic structure package for quantum computers”. *Quantum Science and Technology* **5**, 034014 (2020).
- [51] P. Jordan and E. Wigner. “Über das Paulische Äquivalenzverbot”. *Zeitschrift für Physik* **47**, 631–651 (1928).
- [52] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. “Quantum computational chemistry”. *Rev. Mod. Phys.* **92**, 015003 (2020).
- [53] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. “Quantum Chemistry in the Age of Quantum Computing”. *Chem. Rev.* **119**, 10856–10915 (2019).
- [54] Yasunari Suzuki, Yoshiaki Kawase, Yuya Masumura, Yuria Hiraga, Masahiro Nakadai, Jiabao Chen, Ken M. Nakanishi, Kosuke Mitarai, Ryosuke Imai, Shiro Tamiya, Takahiro Yamamoto, Tennin Yan, Toru Kawakubo, Yuya O. Nakagawa, Yohei Ibe, Youyuan Zhang, Hirotugu Yamashita, Hikaru Yoshimura, Akihiro Hayashi, and Keisuke Fujii. “Qulacs: a fast and versatile quantum circuit simulator for research purpose”. *Quantum* **5**, 559 (2021).
- [55] Zhendong Li, Junhao Li, Nikesh S. Dattani, C. J. Umrigar, and Garnet Kin-Lic Chan. “The electronic complexity of the ground-state of the fmo cofactor of nitrogenase as relevant to quantum simulations”. *J. Chem. Phys.* **150**, 024302 (2019).
- [56] Craig Gidney and Austin G. Fowler. “Efficient magic state factories with a catalyzed $|CCZ\rangle$ to $2|T\rangle$ transformation”. *Quantum* **3**, 135 (2019).
- [57] Seiseki Akibue, Go Kato, and Seiichiro Tani. “Probabilistic unitary synthesis with optimal accuracy”. *ACM Transactions on Quantum Computing* (2024).
- [58] Nobuyuki Yoshioka, Seiseki Akibue, Hayata Morisaki, Kento Tsubouchi, and Yasunari Suzuki. “Error crafting in probabilistic quantum gate synthesis” (2024). [arXiv:2405.15565](https://arxiv.org/abs/2405.15565).

A “Variance” of approximation error

In this section, to justify the use of the average effect derived in Sec. 4, we compute “variance” of this random effect. In this subsection, let $\mathbb{E}$ denote the average within the constraint Eq. (5). Now, we compute a “variance” operator

$$\mathbb{E}[(R_{\mathbf{n}}(\varphi)AR_{\mathbf{n}}^\dagger(\varphi) - \mathbb{E}[R_{\mathbf{n}}(\varphi)AR_{\mathbf{n}}^\dagger(\varphi)])(R_{\mathbf{n}}(\varphi)AR_{\mathbf{n}}^\dagger(\varphi) - \mathbb{E}[R_{\mathbf{n}}(\varphi)AR_{\mathbf{n}}^\dagger(\varphi)])^\dagger]$$

Since

$$\mathbb{E}[R_{\mathbf{n}}(\varphi)AR_{\mathbf{n}}^\dagger(\varphi)] = (1-p)A + p \operatorname{Tr}(A) \frac{I}{2}$$

as we derived above, the “variance” operator is

$$\begin{aligned} & \mathbb{E} \left[\left(R_{\mathbf{n}}(\varphi)AR_{\mathbf{n}}^\dagger(\varphi) - (1-p)A + p \operatorname{Tr}(A) \frac{I}{2} \right)^\dagger \left(R_{\mathbf{n}}(\varphi)AR_{\mathbf{n}}^\dagger(\varphi) - (1-p)A + p \operatorname{Tr}(A) \frac{I}{2} \right) \right] \\ &= \mathbb{E} \left[R_{\mathbf{n}}(\varphi)A^\dagger AR_{\mathbf{n}}^\dagger(\varphi) \right] - \left((1-p)A + p \operatorname{Tr}(A) \frac{I}{2} \right)^\dagger \left((1-p)A + p \operatorname{Tr}(A) \frac{I}{2} \right) \\ &= (1-p)A^\dagger A + p \operatorname{Tr}(A^\dagger A) \frac{I}{2} \\ &\quad - \left((1-p)^2 A^\dagger A + \frac{p(1-p)}{2} (\operatorname{Tr}(A^\dagger)A + \operatorname{Tr}(A)A^\dagger) + p^2 |\operatorname{Tr}(A)|^2 \frac{I}{4} \right) \\ &= p \left((1-p)A^\dagger A - \frac{(1-p)}{2} (\operatorname{Tr}(A^\dagger)A + \operatorname{Tr}(A)A^\dagger) + \left(\operatorname{Tr}(A^\dagger A) - \frac{p|\operatorname{Tr}(A)|^2}{2} \right) \frac{I}{2} \right). \end{aligned}$$

The trace norm of this operator is

$$p \left((2-p) \operatorname{Tr}(A^\dagger A) + \left(\frac{p}{2} - 1 \right) |\operatorname{Tr}(A)|^2 \right) = O(p) = O(\epsilon^2).$$

In this sense, the variance of the approximation error is small when each gate-approximation error ϵ is sufficiently small.

B Unitary coupled-cluster singles and doubles ansatz

In the numerical simulations in the main text, we employed unitary coupled-cluster singles and doubles (UCCSD) ansatz [33]. The UCCSD ansatz is defined by

$$\hat{U}_{\text{UCCSD}}(\boldsymbol{\theta}) = e^{\hat{T}_s - \hat{T}_s^\dagger} e^{\hat{T}_{d1} - \hat{T}_{d1}^\dagger} e^{\hat{T}_{d2} - \hat{T}_{d2}^\dagger}, \quad (22)$$

$$\hat{T}_s = \sum_{o,v} \theta_{ov}^{(s)} \left(\hat{a}_{v\uparrow}^\dagger \hat{a}_{o\uparrow} + \hat{a}_{v\downarrow}^\dagger \hat{a}_{o\downarrow} \right), \quad (23)$$

$$\hat{T}_{d1} = \sum_{o,v} \theta_{ov}^{(d1)} \hat{a}_{v\uparrow}^\dagger \hat{a}_{o\uparrow} \hat{a}_{v\downarrow}^\dagger \hat{a}_{o\downarrow}, \quad (24)$$

$$\hat{T}_{d2} = \sum_{(o_1, v_1) \neq (o_2, v_2)} \sum_{\sigma, \tau = \uparrow, \downarrow} \theta_{o_1 o_2 v_1 v_2}^{(d2)} \hat{a}_{v_1 \sigma}^\dagger \hat{a}_{o_1 \sigma} \hat{a}_{v_2 \tau}^\dagger \hat{a}_{o_2 \tau}, \quad (25)$$

where $a_{p\sigma}(a_{p\sigma}^\dagger)$ is an annihilation (creation) operator of electron in an orbital p with spin σ and o, o_1, o_2 (v, v_1, v_2) are the occupied (virtual) orbitals in the Hartree-Fock state. The circuit parameters of this ansatz are $\boldsymbol{\theta} = (\theta_{ov}^{(s)})_{ov} \cup (\theta_{ov}^{(d1)})_{ov} \cup (\theta_{o_1 o_2 v_1 v_2}^{(d2)})_{o_1 o_2 v_1 v_2}$. To transform this ansatz into qubit operators, we performed the Jordan-Wigner transformation [51] for $\hat{T}_s - \hat{T}_s^\dagger, \hat{T}_{d1} - \hat{T}_{d1}^\dagger$ and $\hat{T}_{d2} - \hat{T}_{d2}^\dagger$. For example, if $\hat{T}_s - \hat{T}_s^\dagger$ is transformed into the sum of multiqubit Pauli operators $i \sum_{ov} \theta_{ov}^{(s)} \sum_k \hat{P}_k^{(ov)}$, we (approximately) implement $e^{\hat{T}_s - \hat{T}_s^\dagger}$ by the Trotterization:

$$\prod_{ov} \prod_k \exp[i\theta_{ov}^{(s)} \hat{P}_k^{(ov)}].$$

The operators $e^{\hat{T}_{d1} - \hat{T}_{d1}^\dagger}$ and $e^{\hat{T}_{d2} - \hat{T}_{d2}^\dagger}$ were implemented as qubit operators in the same way.

Note that every multi-qubit Pauli rotation $\exp[i\theta_{ov}^{(s)} \hat{P}_k^{(ov)}]$ can be expressed as a composite of Clifford gates and a single-qubit z -rotation with the same angle θ_{ov} [33]. Therefore, it suffices to give the Clifford+ T decomposition of the corresponding z -rotation to obtain the decomposition of $\exp[i\theta_{ov}^{(s)} \hat{P}_k^{(ov)}]$.

C Results of Numerical Experiments

In this section, we show the whole results of our numerical experiments, which are partly shown in Sec. 5.3. We consider the molecules shown in Table 1.

In Fig. 4, we show the energy curve of each molecule, and in Fig. 5 we show the fidelity curve and indicate the threshold T -count to achieve the relative error $\delta_{\text{rel}} < 10^{-4}$.

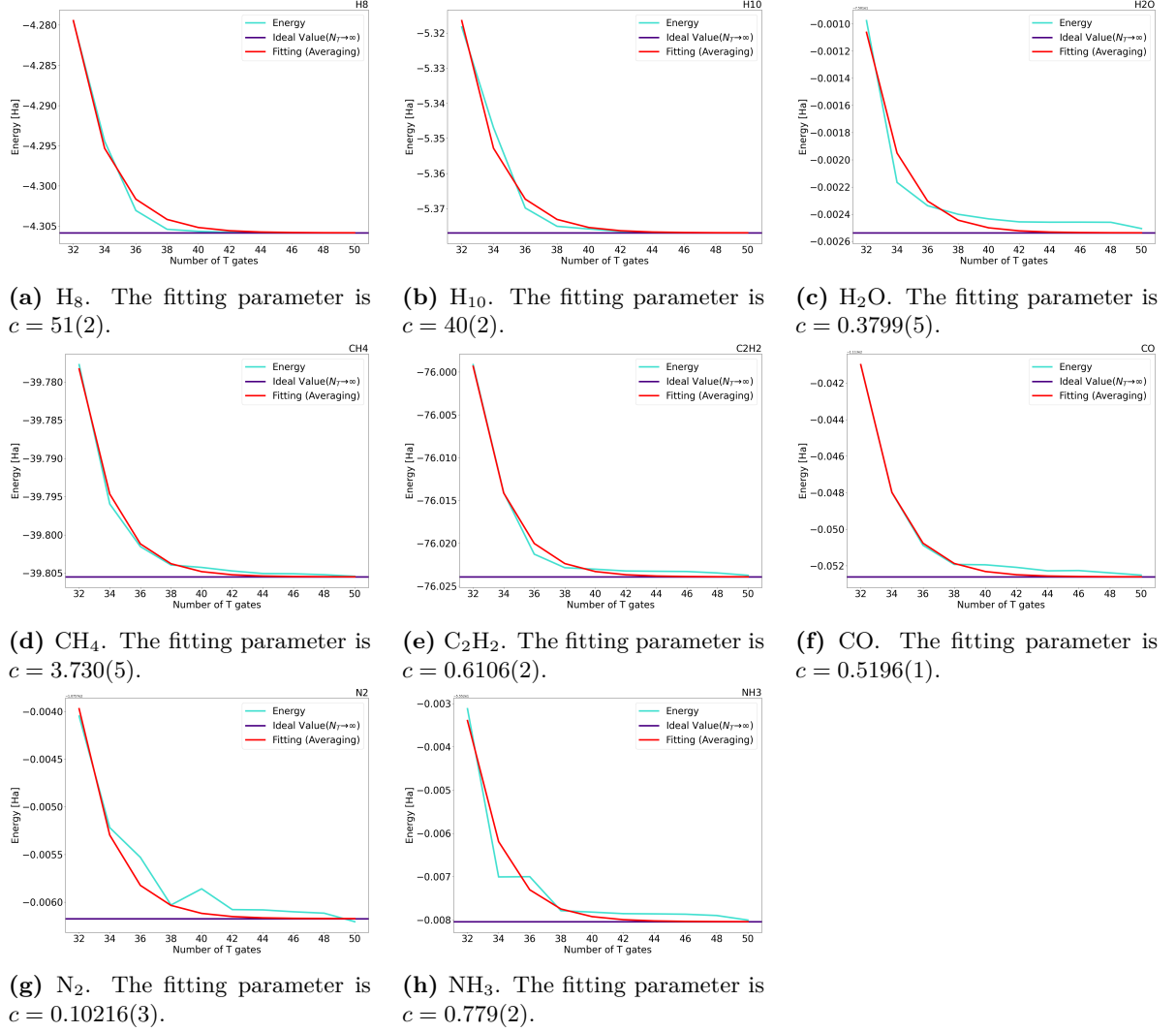
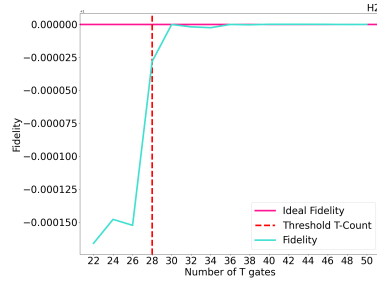
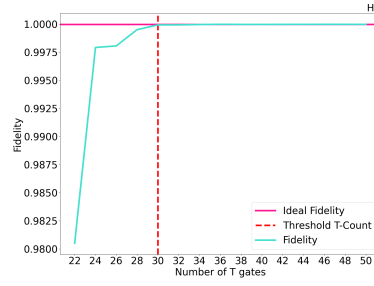


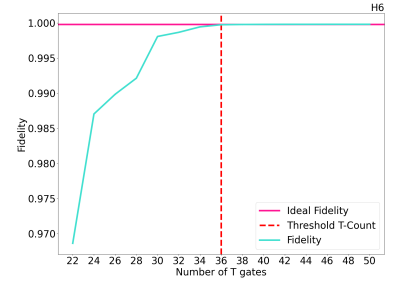
Figure 4: Energy expectation values of Clifford+ T -gate decomposed UCCSD' states with T -gate budget $32 \leq N_T \leq 50$. In each panel, the orange curve represents the expectation values of the UCCSD' state that are decomposed by Algorithm 1. The horizontal purple line indicates the true UCCSD' state energy obtained by the UCCSD ansatz without Clifford+ T decomposition. The blue curve is the fitting of the orange curve modeled by Eq. (20).



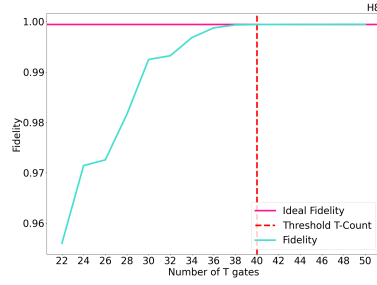
(a) H_2



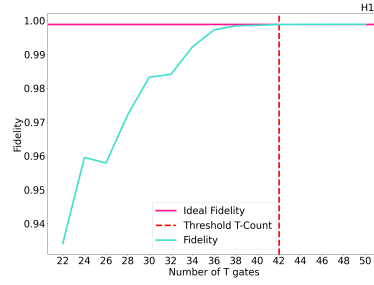
(b) H_4



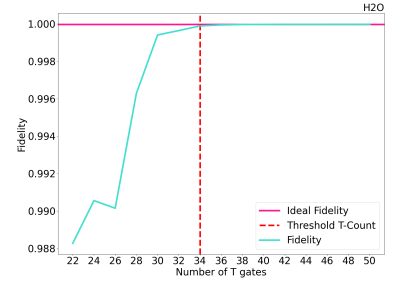
(c) H_6



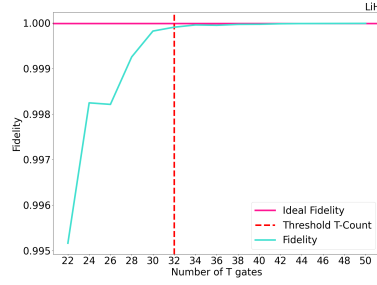
(d) H_8



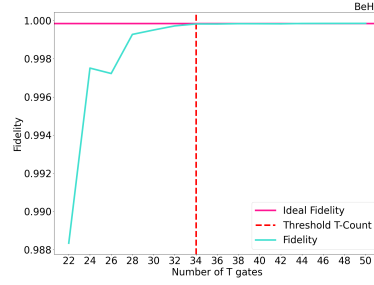
(e) H_{10}



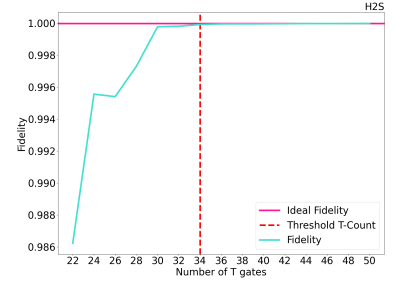
(f) H_2O



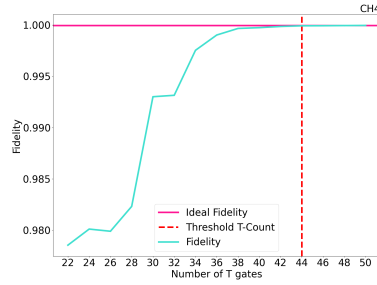
(g) LiH



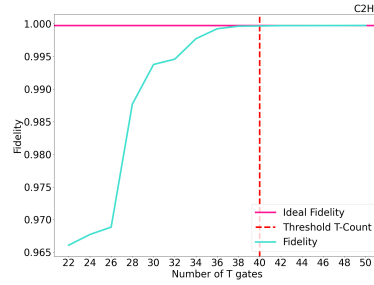
(h) BeH_2



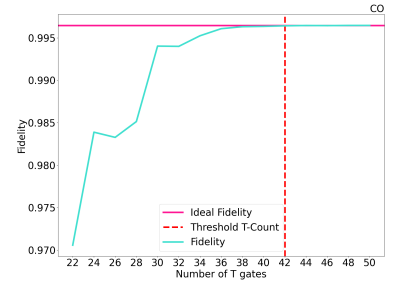
(i) H_2S



(j) CH_4



(k) C_2H_2



(l) CO

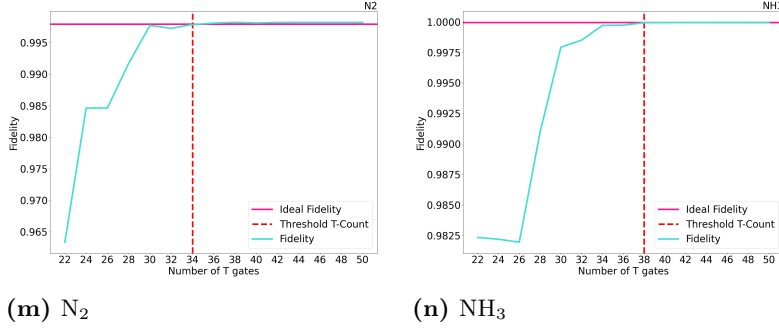


Figure 5: The fidelity between the exact ground state of the Hamiltonian and the Clifford+ T -gate decomposed UCCSD' states with T -gate budget $22 \leq N_T \leq 50$. In each panel, the blue curve represents the fidelity between the exact ground state and the UCCSD' state that is decomposed by Algorithm 1. The horizontal pink line indicates the fidelity between the exact ground state and the true UCCSD' state obtained by the UCCSD ansatz without Clifford+ T decomposition. The orange vertical line represents the T -count threshold that achieves Eq. (21).