Quantum computing enhanced computational catalysis

Journal Article

Author(s):
von Burg, Vera (iD); Low, Guang Hao; Häner, Thomas; Steiger, Damian S.; Reiher, Markus (iD); Roetteler, Martin; Troyer, Matthias (iD)

Quantum computing enhanced computational catalysis

Vera von Burg [1] Guang Hao Low,[2] Thomas Häner [3] Damian S. Steiger [3] Markus Reiher [1,*]
Martin Roetteler,[2] and Matthias Troyer [2,†]

[1]*Laboratorium für Physikalische Chemie, ETH Zürich, Vladimir-Prelog-Weg 2, 8093 Zürich, Switzerland*
[2]*Microsoft Quantum, Redmond, Washington 98052, USA*
[3]*Microsoft Quantum, 8038 Zürich, Switzerland*

The quantum computation of electronic energies can break the curse of dimensionality that plagues many-particle quantum mechanics. It is for this reason that a universal quantum computer has the potential to fundamentally change computational chemistry and materials science, areas in which strong electron correlations present severe hurdles for traditional electronic structure methods. Here we present a state-of-the-art analysis of accurate energy measurements on a quantum computer for computational catalysis, using improved quantum algorithms with more than an order of magnitude improvement over the best previous algorithms. As a prototypical example of local catalytic chemical reactivity we consider the case of a ruthenium catalyst that can bind, activate, and transform carbon dioxide to the high-value chemical methanol. We aim at accurate resource estimates for the quantum computing steps required for assessing the electronic energy of key intermediates and transition states of its catalytic cycle. In particular, we present quantum algorithms for double-factorized representations of the four-index integrals that can significantly reduce the computational cost over previous algorithms, and we discuss the challenges of increasing active space sizes to accurately deal with dynamical correlations. We address the requirements for future quantum hardware in order to make a universal quantum computer a successful and reliable tool for quantum computing enhanced computational materials science and chemistry, and identify open questions for further research.

I. INTRODUCTION

Quantum computing [1–4] has the potential to efficiently solve some computational problems that are exponentially hard to solve on classical computers. Among these problems, one of the most prominent cases is the calculation of quantum electronic energies in molecular systems [5–9]. Due to its many applications in chemistry and materials science, this problem is widely regarded as the "killer application" of future quantum computers [10], a view that was supported by our first rigorous resource estimate study for the accurate calculation of electronic energies of a challenging chemical problem [11].

At the heart of chemistry is predicting the outcome of chemical processes in order to produce chemicals, drugs, or functional molecular assemblies and materials. A prerequisite for this ability to predict chemical processes is an understanding of the underlying reaction mechanisms. Quantum mechanics allows one to assign energies to molecular structures so that a comparison of these energies in a sequence of molecular transformations can be taken as a measure to rate the viability of such a chemical reaction. Relative energies are a direct means to predict reaction heats and activation barriers for chemical processes. However, the reliability of such predictions depends crucially on the accuracy of the underlying energies, of which the electronic energy is often the most important ingredient. This energetic contribution of the dynamics of the electrons in a molecule can be calculated by solving the electronic Schrödinger equation, typically done in a so-called one-particle basis, the set of molecular orbitals.

The computational complexity of an exact solution of the electronic Schrödinger equation on classical computers is prohibitive as the many-electron basis expansion of the quantum state of interest grows exponentially with the number of molecular orbitals (often called the "curse of dimensionality"). An exact diagonalization of the electronic Hamiltonian in the full many-electron representation is therefore hard and limited to small molecules that can be described by comparatively few (on the order of 20) orbitals. Once accomplished, it is said that a full configuration interaction (full-CI) solution in this finite orbital basis has been found. Since typical molecular systems will require on the order of 1000 molecular orbitals for their reliable description, exact-diagonalization methods need to restrict the orbital space to about 20 orbitals chosen from the valence orbital space [called complete active space (CAS) CI]. Accordingly, approximate methods have been

developed in quantum chemistry for large orbital spaces. By contrast, quantum computing allows for an encoding of a quantum state in a number of qubits that scales only linearly with the number of molecular orbitals and has therefore the potential to deliver full-CI solutions for large orbital spaces that are inaccessible to traditional computing because of the exponential scaling.

In our previous case study on the feasibility of quantum computing for chemical reactivity [11], we chose to investigate electronic structures of a biocatalyst with a still unknown mode of action, i.e., the active site of nitrogenase which is a polynuclear iron-sulfur cluster. This choice was motivated by the fact that such polynuclear $3d$ transition metal clusters are known to exhibit strongly correlated electronic structures that are hard to describe reliably with approximate electronic structure methods on classical computers (electronically excited states of molecules in photophysical and photochemical applications represent another example of this class of problems). In Ref. [11], which relied on a structure model of the active site of nitrogenase in its resting state, we avoided biasing toward the electronic ground state of that particular structure by optimizing arbitrary electronic states for that structure for spin and charge states. These did not coincide with the electronic ground state of the resting state as emphasized in the supporting information of Ref. [11]. Although the resource estimates provided in Ref. [11] supported the view that quantum computing is a true competitor of state-of-the-art CAS-CI-type electronic structure methods such as the density matrix renormalization group (DMRG) [12,13] or full configuration interaction quantum Monte Carlo (FCIQMC) [14] for studies of chemical reaction mechanisms, we did not consider an actual reaction mechanism.

In this work we therefore revisit the problem of quantum computing enhanced reaction mechanism elucidation by considering the latest algorithmic advances, but now with a focus on a specific chemical reaction that is prototypical for homogeneous catalysis and equally well relevant for heterogeneous catalysis. The example that we chose is the catalytic functionalization of carbon dioxide, i.e., the capture of the small green-house gas carbon dioxide by a catalyst that activates and eventually transforms it to a useful chemical such as methanol. Hence, we continue to focus on small-molecule activation catalysis, but emphasize that, despite the obvious interest into this specific system in the context of carbon capture technologies, our analysis is of general value to a huge body of chemical reactivity studies. Moreover, we reexamine some of our initial assumptions that are a moving target in the fast developing field of quantum computing. These are the gate counts for the quantum algorithm that performs the energy measurement and assumptions on the error corrected gate times on a future quantum computer, which are heavily dependent on the underlying technology of its hardware. Furthermore, we extend our previous work with respect to the steps that need to be carried out by a quantum computer, specifically regarding the resources of the state preparation step.

Since our previous work [11], there has been significant progress in quantum algorithms, but also a better understanding of what it takes to build a scalable quantum computer has been reached. On the algorithmic side, Hamiltonians represented by a linear combination of unitaries in a so-called black-box query model can now be simulated with optimal cost using techniques called "qubitization" [15] and "quantum signal processing" [16]. In addition, structure in broad families of Hamiltonians can be exploited to reduce the cost of simulation even more. This includes Hamiltonians with geometrically local interactions [17] and large separations in energy scales [18]. Even more recently, very promising tight theoretical bounds on the performance of traditional Lie-Trotter-Suzuki formulas on average-case Hamiltonians have been obtained [19].

Here we introduce further refinements to the technique of qubitization applied to molecular systems, which has already been noted [20] to be particularly promising in terms of the dominant quantum Toffoli-gate complexity [21]. We assume that the two-electron tensor describing interactions between N molecular orbitals has a low-rank approximation in a so-called double-factorized representation. This leads to Toffoli gate cost estimates that are orders of magnitude better. For instance, we previously estimated that obtaining an energy level of a nitrogen fixation problem to 1 mHartree [11] cost the equivalent of 1.5×10^{14} Toffoli gates. This was reduced to 2.3×10^{11} Toffoli gates by Berry *et al.* [22] using a so-called single-factorized representation in the qubitization approach. Under similar assumptions on the rank as Berry *et al.*, we achieve 1.2×10^{10} Toffoli gates, a further improvement of more than an order of magnitude.

We have also taken into account realistic assumptions for mid-term quantum hardware. While our previous work [11] focused on optimistic future devices with logical gate times of 100 ns and all-to-all connectivity between qubits, we here employ gate times of 10 μs for fault tolerant gates with nearest neighbor connectivity—realistic assumptions for mid-term quantum computer architectures. We therefore also discuss the overhead due to mapping the quantum algorithm to a two-dimensional planar layout, which further increases the overall runtime but makes the estimates more realistic.

II. A HOMOGENEOUS CARBON DIOXIDE FIXATION CATALYST

The catalytic process that we selected for our present work is the binding and transformation of carbon dioxide. The infrared absorption properties of carbon dioxide make it a green-house gas that is a major contributor to climate change. Naturally, limiting or even inverting rising carbon dioxide levels in the atmosphere is a truly important goal and all possible ways to accomplish it must be considered. One option, although currently not the most relevant one [23], is carbon dioxide utilization by chemical transformation. In basic research, options have been explored to fix and react inert carbon dioxide to yield chemicals of higher value (see Refs. [24,25] for two recent examples). Also, homogeneous transition metal catalysts have been designed in the past decade, many of them based on ruthenium as the central metal atom. Despite the fact that eventually heterogeneous catalysis may be preferred over homogeneous catalysis, understanding the basics of carbon dioxide fixation chemistry is facilitated by well-defined homogeneous systems. While formate is often the product of such a fixation process,
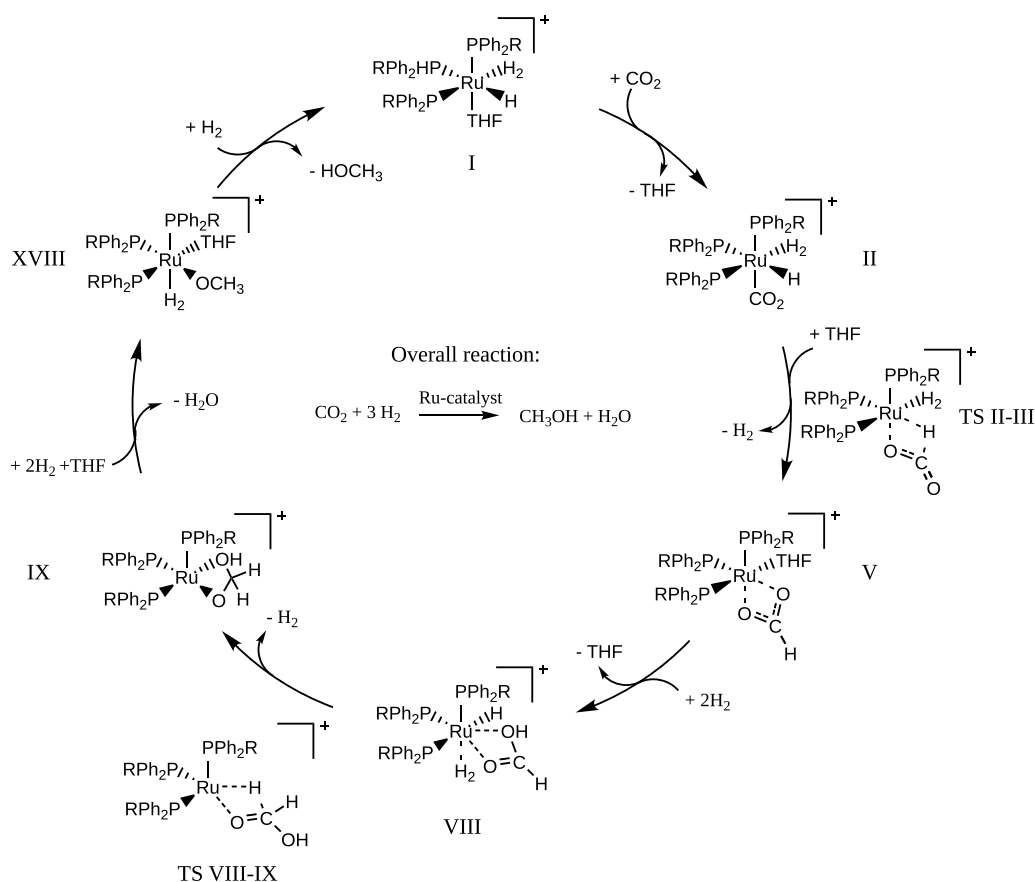
FIG. 1. Selected steps of the catalytic cycle elucidated in Ref. [27] on the basis of DFT calculations: intermediates and transition state structures considered for the present work are shown (Roman numerals are according to the original publication).

methanol is a chemical of higher value. Not many catalysts have been reported so far that can transform carbon dioxide directly into methanol [26] and all of these are plagued by a comparatively low turnover number. To find synthetic catalysts that robustly produce high-value chemicals from carbon dioxide with high turnover number is therefore an important design challenge and computational catalysis can provide decisive insights as well as virtual screening to meet this challenge.

For our assessment of quantum computing resource estimates, we chose a catalyst reported by the Leitner group [27] as extensive density functional theory (DFT) calculations on its mechanism have already been reported by this group. Hence, key molecular structures have already been identified based on DFT. A detailed mechanistic picture has emerged, from which we took intermediates and transition state structures for our resource estimate analysis. Lewis structures of the eight structures selected for our work are shown in Fig. 1.

In this work we focus on the electronic structure of isolated complex structures and therefore neglect any surrounding such as a solvent as well as energy contributions from nuclear dynamics that would be required for the calculation of free energies (cf. Ref. [11] on how to include them). The heavy element ruthenium is known to form complexes that are often low spin, i.e., singlet or doublet, and do not feature strong multiconfigurational character (by contrast to its lighter ho-

molog iron). It is therefore not surprising that we found no pronounced multiconfigurational character through an analysis of occupation numbers as well as orbital entanglement and pair-orbital mutual information measures. Hence, the Ru complexes selected for our resource estimate study feature mostly dynamic electron correlation indicated by small weights of all but one electronic configurations (Slater determinants) that contribute to the exact wave function in a full configuration interaction expansion. For our resource analysis, this is of little importance but highlights the importance of a larger and faster universal quantum computer with a few thousand logical qubits for complete state representation, to accurately capture all relevant dynamical correlations.

The presentation of the CO_2-fixating ruthenium catalyst of the Leitner group [27] was accompanied by extensive DFT calculations of the potentially relevant molecular structures, which is a routine procedure in chemistry. However, DFT electronic energies are plagued by approximations made to the exchange-correlation functional [29,30] and can therefore be of unknown reliability (cf., e.g., Refs. [31–33]). Contemporary DFT is often considered to be a reliable computational approach, but its actual accuracy is in general not known for a reaction under investigation. While molecular equilibrium structures are predicted with sufficient accuracy, the assigned electronic energies can still be affected by significant errors. This can be understood from error measures such as the largest and the mean absolute deviation
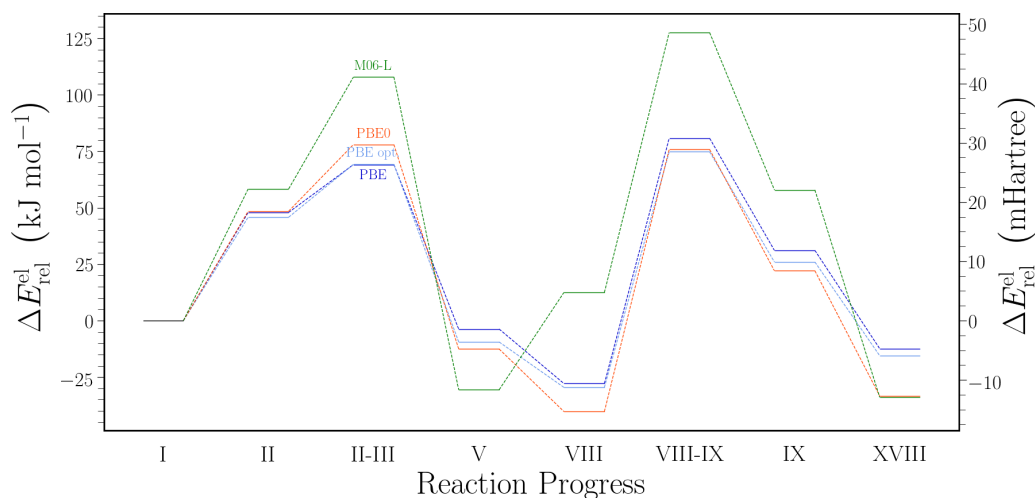
FIG. 2. Relative DFT electronic reaction energies for the catalytic cycle obtained with a def2-TZVP basis set and three different approximate density functionals: M06-L [27], PBE, and PBE0 on M06-L/def2-SVP-optimized structures taken from Ref. [27]. For comparison, we provide relative electronic reaction energies from PBE/def2-TZVP//PBE/def2-TZVP calculations (in Pople's double slash notation, where the density functional behind the double slash is the one for which the structure was obtained) labeled as "PBE opt."

from energy reference data of large balanced benchmark data sets. Another indication is the comparatively poor performance of tailored system-focused parametrizations (see, e.g., Refs. [32,34]). The authors of Ref. [27] selected the Minnesota functional M06-L [35] because they found good overall agreement with their experimental results. In order to demonstrate the uncertainty that generally can affect DFT results, which can be very cumbersome if experimental data are not available for comparison, we supplemented the DFT data of Ref. [27] with results obtained with the standard generalized-gradient-approximation functional PBE [36] and its hybrid variant PBE0 [37] for the eight structures of the reaction mechanism considered in this work (data shown in Fig. 2).

As can be seen in Fig. 2, whereas structure optimization (compare results for structures optimized with the same functional with the single-point data denoted by the double slash notation) has a small effect on the relative electronic energies, the difference between the functionals can amount to more than 50 kJ/mol (i.e., 19 mHartree per molecule) and can even reverse the qualitative ordering of the compounds (compare structures V and VIII). The variation of the electronic DFT energy obtained with different density functional approximations is an indication of the magnitude by which DFT energies can change with respect to this key approximation in DFT. However, we want to emphasize that one does not know to what extent any of these energies for any of these structures is accurate. This can lead to qualitatively, but also quantitatively wrong conclusions about the mechanism (recall the fact that microkinetic models will require errors in relative energies to be smaller than about 1 mHartree as they enter rate expressions through an error-amplifying exponential expression; see Ref. [34]). Obviously, in the absence of any additional information (such as experimental data or more accurate calculations), it is virtually impossible to settle on reliable energetics with potentially dramatic consequences for the elucidation of a reaction mechanism.

III. QUANTUM COMPUTING ENHANCED COMPUTATIONAL CATALYSIS

The elucidation of chemical reaction mechanisms is routinely accomplished with approximate quantum chemical methods [38,39], for which usually stationary structures on Born-Oppenheimer potential energy surfaces are optimized, yielding stable reactants, products, and intermediates of a chemical process. More importantly, optimized first-order saddle points on such a surface represent transition state structures, which are key for detailed kinetic studies and hence for the prediction of concentration fluxes.

In the following we discuss where in the mechanism elucidation process quantum computing can be efficient, useful, and decisive, and hence, how quantum computing can efficiently enhance and reinforce computational catalysis to make a difference. As there are many steps involved that require a deep understanding of various branches of theoretical chemistry, we provide an overview of the essential steps in Fig. 3. Understanding chemical catalysis, and chemical reactions in general, requires an exploration of relevant molecular structures (*structure exploration* in Fig. 3), which is usually done manually and with DFT approaches (as in Ref. [27] for our example here), but which can now also be done in a fully automated and even autonomous way [40]. These structures need to be assigned an energy which may be conveniently separated into an electronic contribution (steps 1–7 in Fig. 3) and a remaining part (*additional free energy calculations* in Fig. 3) containing nuclear and other effects (calculated in a standard rigid-rotor–harmonic-oscillator model accompanied by dielectric continuum embedding or by explicit molecular dynamics) to eventually yield a free energy calculated from all related microstates. Relative free-energy differences will eventually be used as barrier heights in expressions for absolute rate constants (*kinetic modeling* in Fig. 3) that can then be used in kinetic modeling for predicting concentration fluxes through the chemical web of relevant molecular structures. Ultimately, such knowledge can be exploited to improve on
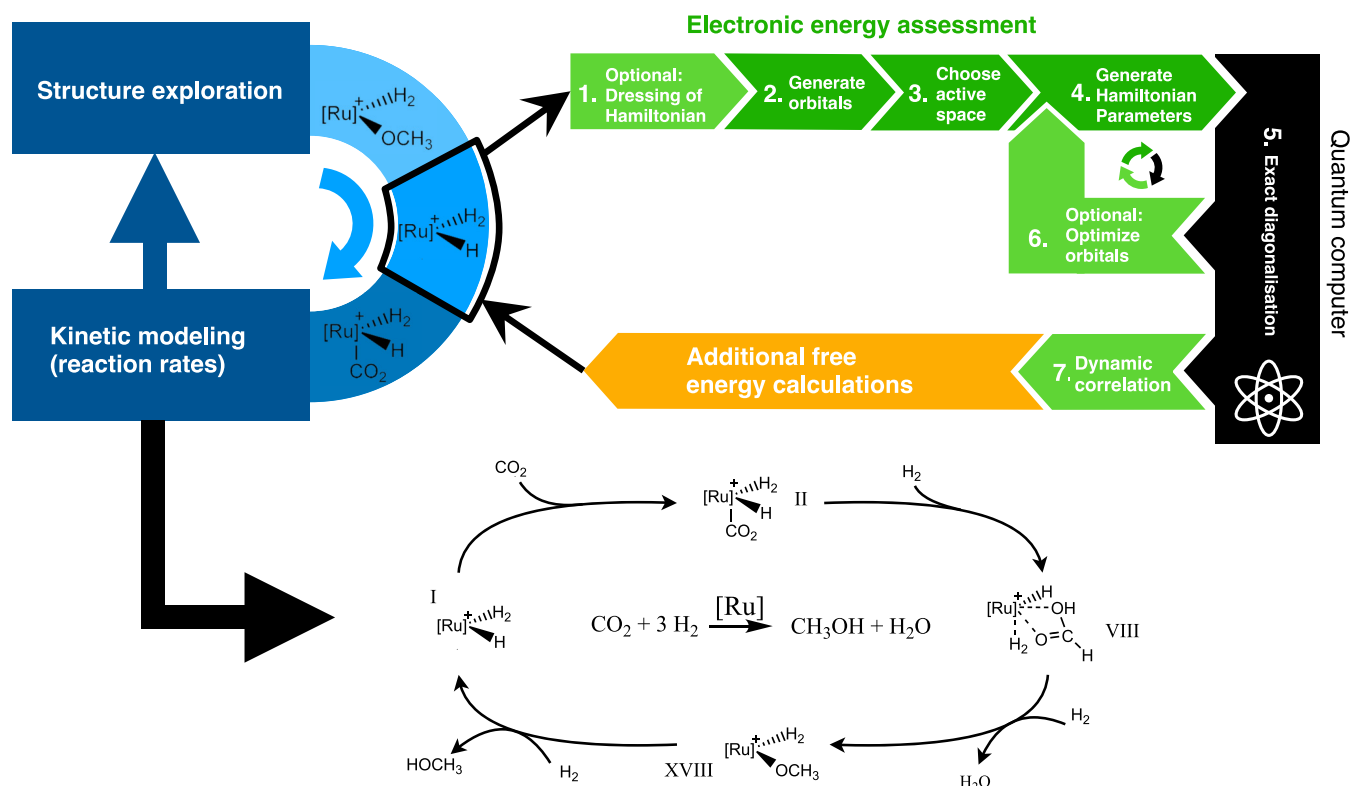
FIG. 3. Protocol of computational catalysis with the key step of quantum computing embedded in black, which is usually accomplished with traditional methods such as CASSCF, DMRG, or FCIQMC (see text for further explanation).

existing or to design new catalysts with enhanced catalytic properties.

Accuracy matters: A reaction rate depends exponentially on the energy difference between a transition state structure and its corresponding stable reactants, which are connected by an elementary reaction step. Because of this exponential dependence, highly accurate energy differences are decisive. While many contributions enter these free energy differences, the electronic energy difference is the most crucial one in bond breaking and bond making processes.

Electronic energies are key components (steps 1–7 in Fig. 3): Electronic energies are notoriously difficult to calculate and standard approximations are affected by unknown errors that can be large. Only for electronically simple structures, so-called closed-shell single-determinant electronic structures, well-established methods exist that run efficiently on a classical computer (such as explicitly correlated local coupled cluster schemes with, at least, perturbatively treated triple excitations [41]). For general electronic structures, however, no method of comparable accuracy exists that is at the same time feasible for moderately sized molecules. In particular, for strong correlation cases, which require more than one Slater determinant for a qualitatively correct description of the electronic wave function, it can be hard to obtain an accurate total electronic energy, which then enters the calculation of relative energies.

In our previous work [11] we considered a quantum computer of moderate size within reach in the not too distant future. Moreover, we assumed that such a machine might have 100 to 200 logical qubits available for the representation of a quantum state. Such a state would be constructed from single-particle states, i.e., molecular orbitals for molecular structures. For decades it has been the goal of quantum chemistry to devise methods that efficiently construct approximations to a many-electron state represented in terms of orbitals. If the full many-electron Hamiltonian is expressed and diagonalized in a complete many-electron (determinantal) basis constructed from such a one-electron basis, then a full-CI calculation may be carried out. Such an exact diagonalization is, however, routinely only feasible for about 18 orbitals [42] (a record calculation was recently carried out for 24 orbitals [43]) due to the exponentially growing number of many-electron states with the number of orbitals.

Required accuracy: A reasonable target accuracy for relative energies (and therefore also for total electronic energies of individual molecular species) is about 1 mHartree, if not 0.1 mHartree. This corresponds to about 2.6 and 0.26 kJ/mol, respectively. Note that thermal energy RT (T being temperature and R the gas constant) at room temperature is on the order of 2.6 kJ/mol, which may be related to the kinetic energy of a reactant molecule at average velocity (compare this to the spread observed for different DFT functionals in our example in the last section).

It is important to understand that this target accuracy is important for relative energies, i.e., for the energy differences that eventually enter the rate constant expressions. For total electronic energies, however, this accuracy does not match the precision with which these energies are actually known. In fact, the true total energies are typically off by a huge energy offset because one does not make an effort to describe core electrons, which contribute significantly to the total electronic energy, but not to reaction chemistry as they are atomically conserved. Hence, such calculations rely on significant error cancellation effects that occur when atomic contributions to the total electronic energy drop out in the calculation of reaction energies (which are relative energies) as they are conserved during a reaction (consider, e.g., a molecular orbital that is mostly of $1s$-atomic orbital character and remains unaffected by a chemical reaction, but contributes significantly to the total electronic energy).

Challenges of electronic structure (steps 1–7 in Fig. 3): It is therefore most important to get the electronic (valence) structure of each relevant molecular structure right. Here quantum computing offers an opportunity [11]. It is important to realize that a typical chemical catalysis problem does not involve very many valence orbitals in every elementary reaction step. As a consequence, the size of the active orbital space, from which the electronic wave function is constructed, does usually not need to be very large and can be easily handled with methods such as complete active space self-consistent field (CASSCF), DMRG, or FCIQMC. The latter two are capable of handling active orbital spaces of up to about 100 spatial orbitals owing to efficient approximations. However, sometimes there may be a price to pay for these approximations and that is a residual uncertainty regarding convergence of the electronic energy. For instance, a DMRG result will critically depend on—apart from a fixed finite value of the bond dimension—proper convergence of the sweeping algorithm, which, at times, might be difficult to determine. In the case of the more recent FCIQMC approach, which is under continuous development and offers extraordinarily efficient scaling on large traditional parallel computers, convergence with respect to the number of walkers or extrapolation to an infinite number of walkers may be hard to achieve for certain molecules. We also note that it is generally hard for any quantum chemical approach (hence, also for DMRG and FCIQMC) to deliver rigorous yet useful information about the error associated with a specific result. Moreover, any active-space approach is plagued by a severe drawback discussed below, namely the neglect of dynamic electronic correlation arising from the majority of virtual orbitals neglected.

Importance of dynamic electron correlation (steps 1 and 7 in Fig. 3): As noted already above, moderately sized molecules, such as the ones studied in this work, may easily require on the order of 1000 or more spatial molecular orbitals for a description that may be considered accurate within the chemically relevant accuracy of about 1 mHartree or even 0.1 mHartree for relative energies. However, the restriction to the valence orbital space from which the active space of the most strongly correlated orbitals is chosen [44,45] compromises this accuracy (note that these active orbitals may be identified based on natural-orbital occupation numbers [46] or on orbital entanglement measures [47,48], even in a completely auto-

mated fashion [47,49]). The vast majority of orbitals that are weakly correlated and neglected in this procedure give rise to dynamic electron correlations, which are neglected in a small-CAS calculation. However, the dynamic electron correlation contribution to the total electronic energy is decisive and so standard recipes exist to approximate it. The most prominent one is the *a posteriori* correction (step 7 in Fig. 3) provided by multireference perturbation theory [50], which, however, requires elements of the three- and four-body reduced density matrices (3-RDMs and 4-RDMs, respectively) that would be extremely hard to obtain by quantum computing. Not even approximate approaches that rely on at most some 3-RDM elements will be accessible by quantum computing for interesting molecules. To evade such a computational bottleneck, perturb-then-diagonalize approaches (step 1 in Fig. 3) such as range-separated DFT for CAS-type methods [51–53] or transcorrelation approaches [54] were proposed for quantum computing [11].

Quantum computing is supposed to be a valuable and, in the long run, an ultimately superior competitor to the aforementioned traditional methods (i) because rigorous error estimates are available and (ii) because systems may be accessible that are traditionally not feasible because of the curse of dimensionality when a total state is to be represented in a large set of orbitals. Its true benefits will fully unfold if energy measurements in the full orbital basis become feasible (on this, see the discussion in the Conclusions section).

Four-index transformation (step 4 in Fig. 3): A cumbersome technical step, to be carried out by traditional computing on classical computers, is the four-index transformation of the two-electron interaction integrals in the Hamiltonian, which scales with the fifth power of the number of basis functions. In this transformation, the final parameters for the electronic Coulomb Hamiltonian in the molecular orbital basis are produced from the four-index integrals defined in the atomic orbital basis, i.e., in the basis which is provided for the representation of all molecular orbitals (typically a set of Gaussian-type functions such as the def2-TZVP basis set used in the DFT calculations presented above). Note, however, that this transformation, while being expensive but feasible for a small CAS, becomes a threat to the whole calculation when the active space grows to eventually incorporate the complete one-electron (atomic orbital) basis (of say, more than 1000 basis functions). Hence, the sheer number of two-electron parameters in the electronic Hamiltonian will require us to rethink how to deal with these terms growing to the fourth power in the number of orbitals, such as in recent work [55] that generates sparse low-rank approximations to the two-electron integrals. Naturally this issue has also been discussed in modern traditional approaches [56].

IV. QUANTUM ALGORITHMS FOR CHEMISTRY

For the catalysis problem we require quantum algorithms that provide reliable results with controlled errors on the electronic energy in a given orbital basis. Uncontrolled approximations in quantum algorithms would negate the advantages offered by quantum computers, which will require tremendous effort to build and operate even at a moderately sized error-corrected scale.

Though it is without doubt that the popular variational quantum eigensolver (VQE) [57] can obtain parameters of a unitary coupled cluster (UCC) parametrization of the electronic wave function that is likely to be accurate (especially when higher than double excitations are considered [58]), this scheme unavoidably generates a residual unknown uncertainty in the true electronic energy. Reducing these errors by improving the ansatz will require significantly more gate operations than are expected to be possible on non-error-corrected noisy intermediate scale quantum devices [59,60]. Moreover, the number of repetitions required to estimate energies with sufficient accuracy of 1 mHartee or better is enormous [61].

We also note that knowledge of reliable and controllable errors in quantum algorithms is a decisive advantage over classical methods such as DMRG and FCIQMC, for which convergence control with respect to their parameters is not necessarily easy or even feasible.

Therefore, we turn to one of the most promising applications of quantum computers, which is a bounded-error simulation of quantum systems using quantum phase estimation. The main idea is to synthesize a quantum circuit that implements the real time-evolution operator $W = e^{-iH/\alpha}$ by a given Hamiltonian H for some normalizing factor α, which henceforth are always in atomic units of Hartree. When applied n times to an eigenstate $H|\psi_k\rangle = E_k|\psi_k\rangle$, a phase nE_k/α is accumulated. Quantum phase estimation then estimates the energy E_k with a standard deviation $\Delta_E = O(\alpha/n)$ [62]. If one prepares an arbitrary trial state $|\psi_{\text{trial}}\rangle$ rather than an eigenstate, phase estimation collapses the trial state to the kth eigenstate with probability $p_k = |\langle\psi_k|\psi_{\text{trial}}\rangle|^2$ and it returns an estimate to the corresponding energy E_k [63].

The phase estimation procedure is executed on a quantum computer by applying a sequence of quantum gates. If the unitary W is implemented using a number c_W of quantum gates, the overall quantum gate cost of obtaining a single estimate $\hat{E}_k$ is then

$$c_W \frac{\pi\alpha}{2\Delta_E},\tag{1}$$

where the factor $\frac{\pi\alpha}{2\Delta_E}$ arises from previous analyses on the performance of phase estimation [20,64] (combined with a so-called phase-doubling trick [20,65]). In general, the quantum circuit only approximates W to some bounded error Δ_W in spectral norm. This adds a systematic bias of $\alpha\Delta_W$ to the estimate $\hat{E}_k$. Thus, we budget for this error by making the somewhat arbitrary choice of performing phase estimation to an error of $0.9\Delta_E$, and compiling W so that $\Delta_W \leqslant 0.1\Delta_E/\alpha$. To date, there are several prominent quantum algorithms for approximating real time evolution, such as Lie-Trotter-Suzuki product formulas [66], sparse Hamiltonian simulation [67], linear combination of unitaries [68], qubitization [15], and quantum signal processing [16]. In this work we consider the electronic Hamiltonian in its nonrelativistic form with Coulomb interactions (in Hartree atomic units),

$$H = \sum_{ij,\sigma} h_{ij} a^\dagger_{(i,\sigma)} a_{(j,\sigma)}$$

$$+ \frac{1}{2} \sum_{ijkl,\sigma\rho} h_{ijkl} a^\dagger_{(i,\sigma)} a^\dagger_{(k,\rho)} a_{(l,\rho)} a_{(j,\sigma)},\tag{2}$$

which is parametrized through the one- and two-electron integrals h_{ij} and h_{ijkl} of the molecular orbitals $\{\psi_i\}$,

$$h_{ij} = \int \psi^*_i(x_1)\left(-\frac{\nabla^2}{2} - \sum_m \frac{Z_m}{|x_1 - r_m|}\right)$$
$$\times \psi_j(x_1)d^3x_1,\tag{3}$$

$$h_{ijlk} = \int \psi^*_i(x_1)\psi_j(x_1)\left(\frac{1}{|x_1 - x_2|}\right)$$
$$\times \psi^*_k(x_2)\psi_l(x_2)d^3x_1 d^3x_2,\tag{4}$$

where x_i denote electronic coordinates and Z_m is the charge number of nucleus m at position r_m (note that a relativistic generalization is straightforward [69]). We explicitly separate the fermion indices $p \equiv (i, \sigma)$ into an index where $i \in \{1, \ldots, N\}$ enumerates the N spatial molecular orbitals, and $\sigma \in \{0, 1\}$ indexes spin up and spin down. Hence, the fermion operators satisfy the usual anticommutation relations

$$\{a_p, a_q\} = 0, \quad \{a^\dagger_p, a^\dagger_q\} = 0, \quad \{a_p, a^\dagger_q\} = \delta_{pq}\mathcal{I},\tag{5}$$

and the coefficients h_{ij}, h_{ijkl} are real and satisfy the symmetries

$$h_{ij} = h_{ji},$$
$$h_{ijkl} = h_{jikl} = h_{ijlk} = h_{jilk} = h_{lkij}$$
$$= h_{lkji} = h_{klij} = h_{klji}.\tag{6}$$

For the purposes of phase estimation through the so-called *qubitization* approach, it is not necessary to simulate time-evolution e^{-iHt}. Whereas our previous work [11] approximated the time-evolution operation using Lie-Trotter-Suzuki product formulas, it can be advantageous in some cases to implement the unitary walk operator $W = e^{i\sin^{-1}(H/\alpha)}$ [15,70,71] instead, which has some normalizing constant $\alpha \geqslant \|H\|$ that ensures the arcsine is real. This walk operator can be implemented exactly, assuming access to arbitrary single-qubit rotations, in contrast to all known quantum algorithms where time-evolution e^{-iHt} can only be approximated. After estimating the phase $\hat{\theta} = \sin^{-1}(E_k/\alpha)$ with phase estimation, we may obtain $\hat{E}_k$ by applying $\sin(\hat{\theta})$ in a classical postprocessing step.

We focus on the qubitization technique applied to the electronic Hamiltonian of Eq. (2), with the goal of minimizing the quantum gate costs in Eq. (1). In fault-tolerant architectures, quantum gate costs reduce to the number of so-called primitive Clifford gates (e.g., Hadamard$= \frac{1}{\sqrt{2}}[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}]$, phase$= [\begin{smallmatrix} 1 & 0 \\ 0 & i \end{smallmatrix}]$, and controlled-NOT) and non-Clifford gates (e.g., $T = [\begin{smallmatrix} 1 & 0 \\ 0 & \sqrt{i} \end{smallmatrix}]$ and Toffoli, which applies a NOT gate controlled on two input bits being in the "one" state. As the physical resources needed to implement a single error-corrected primitive non-Clifford gate such as the T gate are on the order of 100 to 10 000 times higher than for a single two-qubit Clifford gate [72,73]—and one Toffoli gate may be implemented by four T gates—much recent work has focused on optimizing the non-Clifford-gate cost of W. Within an atom-centered basis set (such as the ones employed in this work), this focus has reduced the T gate cost of obtaining a single estimate $\hat{E}_k$ from $\tilde{O}(N^5/\Delta_E^{3/2})$ using Trotter methods to $\tilde{O}(\alpha_{\text{CD}}N^{3/2}/\Delta_E)$ Toffoli gates and $O(N^{3/2}\log N/\Delta_E)$ qubits

using qubitization combined with a so-called single-factorized Hamiltonian representation, where $\alpha_{CD} = O(N^3)$ [22] is a certain norm of the Hamiltonian. Our main technical contribution is a further reduction of the gate cost of W and the normalizing factor to $\alpha_{DF} = O(N^{3/2})$ as described in the next section.

V. EFFICIENT ENCODING OF DOUBLE-FACTORIZED ELECTRONIC STRUCTURE

Our main algorithmic advance is based on a quantum algorithm to "qubitize" the so-called double-factorized representation H_{DF} of the electronic Hamiltonian H. On the one hand, the double-factorized representation is sparse, and therefore minimizes the cost c_W of qubitization, which generally scales with the number of terms needed to represent the Hamiltonian. On the other hand, the double-factorized representation is a partial diagonalization of the original Hamiltonian, and hence has a small normalizing constant α_{DF}. Despite these favorable properties, previous quantum simulations using this representation are based on Trotter methods [74,75].

A. Double-factorized Hamiltonian

The double-factorized form builds upon the single-factorized representation H_{CD} of the Hamiltonian H where

$$H_{CD} \doteq \sum_{ij,\sigma} \tilde{h}_{ij} a^{\dagger}_{(i,\sigma)} a_{(j,\sigma)} + \frac{1}{2} \sum_{r \in [R]} \left(\sum_{ij,\sigma} L^{(r)}_{ij} a^{\dagger}_{(i,\sigma)} a_{(j,\sigma)} \right)^2, \tag{7}$$

$$\tilde{h}_{ij} \doteq h_{ij} - \frac{1}{2} \sum_{l} h_{illj}. \tag{8}$$

Note that the rank-R factorization of the two-electron tensor $h_{ijkl} = \sum_{r \in [R]} L^{(r)}_{ij} L^{(r)\top}_{kl}$ of Eq. (2) into the $N \times N$ symmetric matrices $L^{(r)}$ always exists due to the symmetry constraints of Eq. (6), and may be computed using either a singular-value decomposition or a Cholesky decomposition. This representation also facilitates a low-rank approximation by truncating the rank R. In the worst case, $R \leqslant N^2$. However, it was noted by Peng *et al.* [55] that rank $R \sim N \log N$ for typical molecular systems when N is proportional to the number of atoms, which is a provable statement for 1D systems [76]. This reduces the number of terms needed to describe the second-quantized Hamiltonian from $O(N^4)$ in Eq. (2) to $O(RN^2)$ in Eq. (7). This representation was first exploited by Berry *et al.* [22] to qubitize H_{CD} with a normalizing constant $\alpha_{CD} \doteq 2\|\tilde{h}\|_{EW} + 2 \sum_{r \in [R]} \|L^{(r)}\|^2_{EW}$ expressed using the entry-wise norm $\|L^{(r)}\|_{EW} \doteq \sum_{ij \in [N]} |L^{(r)}_{ij}|$.

The technical innovation in our approach is a quantum circuit, detailed in the Supplemental Material [28], for qubitzing the double-factorized Hamiltonian

$$H_{DF} = \sum_{ij,\sigma} \tilde{h}_{ij} a^{\dagger}_{(i,\sigma)} a_{(j,\sigma)}$$

$$+ \frac{1}{2} \sum_{r \in [R]} \left(\sum_{ij,\sigma} \sum_{\substack{m \\ \in [M^{(r)}]}} \lambda^{(r)}_m \vec{R}^{(r)}_{m,i} \vec{R}^{(r)}_{m,j} a^{\dagger}_{(i,\sigma)} a_{(j,\sigma)} \right)^2. \tag{9}$$

This is obtained by a rank-$M^{(r)}$ eigendecomposition of the symmetric matrices $L^{(r)} = \sum_{m \in [M^{(r)}]} \lambda^{(r)}_m \vec{R}^{(r)}_m \cdot (\vec{R}^{(r)}_m)^{\top}$ into a total of $M \doteq \sum_{r \in [R]} M^{(r)}$, each normalized to be of unit length $\|\vec{R}^{(r)}_m\|_2 = 1$. The number of terms in Eq. (9) is even further reduced to $O(MN)$ as Peng *et al.* [55] also noted that for typical molecular systems where $N \gg 10^3$ scales with the number of atoms, one may retain $M \sim N \log N$ eigenvectors and truncate the rest. A key technical step in our approach is to work in the Majorana representation of fermion operators

$$\gamma_{p,0} = a_p + a^{\dagger}_p, \tag{10}$$

$$\gamma_{p,1} = -i(a_p - a^{\dagger}_p), \tag{11}$$

$$\{\gamma_{p,x}, \gamma_{q,y}\} = 2\delta_{pq}\delta_{xy}\mathcal{I}. \tag{12}$$

As we show in the Supplemental Material, this representation maps

$$H_{DF} \rightarrow \left(\sum_i h_{ii} - \frac{1}{2} \sum_{il} h_{illi} + \frac{1}{2} \sum_{il} h_{llii} \right) \mathcal{I}$$

$$+ \text{One}_{L^{(-1)}} + \frac{1}{2} \sum_{r \in [R]} \text{One}^2_{L^{(r)}}, \tag{13}$$

which is expressed as a sum of squares of one-body Hamiltonians

$$\text{One}_L \doteq \frac{i}{2} \sum_{ij} \sum_{\sigma} L_{ij} \gamma_{i,\sigma,0} \gamma_{j,\sigma,1}, \tag{14}$$

$$L^{(-1)}_{ij} \doteq h_{ij} - \frac{1}{2} \sum_l h_{illj} + \sum_l h_{llij}. \tag{15}$$

This leads to a normalizing constant

$$\alpha_{DF} = \|L^{(-1)}\|_{SC} + \frac{1}{4} \sum_{r \in [R]} \|L^{(r)}\|^2_{SC}. \tag{16}$$

Note that the difference between $L^{(-1)}$ and $\tilde{h}$ is the mean-field Coulomb repulsion contribution term $\sum_l h_{llij}$. Note that α_{DF} is also significantly smaller than α_{CD} from previous approaches [22]. In addition to the factor of 8 reduction in the prefactor of the two-electron norms, the dependence of α_{DF} on Schatten norms $\|L^{(r)}\|_{SC}$ is beneficial as they can be up to a factor of N smaller than the entry-wise norms $\|L^{(r)}\|_{EW}$ that α_{CD} depends on, as follows from the following tight inequalities for any Hermitian $N \times N$ matrix h,

$$\|h\|_{SC} \doteq \sum_{k \in [N]} |\text{Eigenvalues}[h]_k|, \tag{17}$$

$$\frac{1}{N}\|h\|_{EW} \leqslant \|h\|_{SC} \leqslant \|h\|_{EW}, \tag{18}$$

which we prove in the Supplemental Material and may be of independent interest.

The Toffoli gate complexity of synthesizing the walk operator to an error $\Delta_W = 0.1\Delta_E/\alpha_{DF}$, as detailed in the Supplemental Material, is then

$$c_W \leqslant \frac{2M}{1 + \lambda} + 2\lambda N\beta + 8N\beta + 4N$$

$$+ O[\sqrt{R \log M} + \sqrt{M \log (1/\Delta_W)}], \tag{19}$$

for any choice of integer $\lambda \geqslant 0$, and where the parameter $\beta = \lceil 5.652 + \log_2(\frac{N}{\Delta_W}) \rceil$. This also uses the following number of qubits:

$$N\beta(1 + \lambda) + 2N + O[\log{(N/\Delta_W)}]. \qquad (20)$$

Note that λ controls the number of ancillary qubits. By choosing $\lambda = O(\sqrt{\frac{M}{N\beta}})$, the Toffoli cost is $c_W = O(\sqrt{MN\beta} + N\beta)$, which is advantageous when $M \gg N\beta$, using $O(\sqrt{MN\beta} + N + \beta)$ qubits. Assuming the empirical scaling of M and R by Peng *et al.*, our algorithm encodes electronic spectra for atom-centered basis sets into the walk operator using only $c_W = \tilde{O}(N)$ Toffoli gates, and with a normalizing constant α_{DF}, which improves upon the $\tilde{O}(N^{3/2})$ Toffoli gates and larger normalizing constant α_{CD} required by the single-factorized approach [22].

In the following examples we consider typical values of λ that minimize the Toffoli costs between 1 and 5, and the Toffoli gate counts and qubit counts we quote exclude the subdominant big-$O(\cdot)$ component of Eqs. (19) and (20).

B. Truncation

We now describe our procedure for obtaining, by truncating eigenvalues, low-rank approximations $\tilde{H}_{DF}$ of the double-factorized Hamiltonian in Eq. (13) from an initial numerically exact representation H_{DF}. We focus on truncating the two-electron terms as those tend to dominate the cost of block encoding. Given a target approximation error ϵ, we consider two truncations schemes, which we call *coherent* and *incoherent*. The coherent scheme upper bounds the actual error

$$\|\tilde{H}_{DF} - H_{DF}\| \leqslant \epsilon_{co}, \qquad (21)$$

in spectral norm for any given choice of ϵ_{co}. This upper bound is obtained by a triangle inequality, which assumes that all truncated terms have an error that adds linearly. That is, if the difference $\tilde{H}_{DF} - H_{DF} = \sum_j H_j$ is a sum of terms, then we truncate so that

$$\sum_j \|H_j\| \leqslant \epsilon_{co}. \qquad (22)$$

We find that this bound is often quite loose, and the gap between the actual error and ϵ grows with system size. This motivates the *incoherent* scheme, which assumes that the error of truncated terms add incoherently by a sum-of-squares. Thus we truncate so that

$$\sqrt{\sum_j \|H_j\|^2} \leqslant \epsilon_{in}. \qquad (23)$$

Suppose we remove a single eigenvalue $\lambda_m^{(r)}$ from Eq. (13) during truncation. Using the identity $A^2 - (A - B)^2 = AB + BA - B^2$, the difference

$$H_{DF} - \tilde{H}_{DF}$$
$$= \frac{1}{2}\left[\left\{ \text{One}_{L^{(r)}}, \frac{\lambda_m^{(r)}}{2} \sum_\sigma \gamma_{\bar{R}_m^{(r)},\sigma,0} \gamma_{\bar{R}_m^{(r)},\sigma,1} \right\} \right.$$
$$\left. - \left(\frac{\lambda_m^{(r)}}{2} \sum_\sigma \gamma_{\bar{R}_m^{(r)},\sigma,0} \gamma_{\bar{R}_m^{(r)},\sigma,1} \right)^2 \right] \qquad (24)$$

can be bounded in $\|\cdot\|$ as follows:

$$\|H_{DF} - \tilde{H}_{DF}\| \leqslant |\lambda_m^{(r)}| \sum_{n \neq m} |\lambda_n^{(r)}| + \frac{1}{2}|\lambda_m^{(r)}|^2$$
$$= |\lambda_m^{(r)}| \left(\|L^{(r)}\|_{SC} - \frac{1}{2}|\lambda_m^{(r)}| \right)$$
$$\leqslant \|L^{(r)}\|_{SC}|\lambda_m^{(r)}|. \qquad (25)$$

In the coherent scheme we truncate eigenvalues in the index set $\mathcal{T}$ such that $\sum_{(r,m) \in \mathcal{T}} \|L^{(r)}\|_{SC}|\lambda_m^{(r)}| \leqslant \epsilon_{co}$. In the incoherent scheme we truncate such that $\sqrt{\sum_{(r,m) \in \mathcal{T}} (\|L^{(r)}\|_{SC}|\lambda_m^{(r)}|)^2} \leqslant \epsilon_{in}$. In both cases we may maximize the number of truncated eigenvalues by deleting those with the smallest value of $\|L^{(r)}\|_{SC}|\lambda_m^{(r)}|$ first. To evaluate the validity of the truncation schemes, we have computed the error in the DMRG ground-state energies they introduce for a set of benchmark systems (Fig. 4).

In the following we truncate according to the sum-of-square procedure. While this approach does not rigorously bound the total error, we find that it better matches the error we observed for our benchmark systems, but still tends to significantly overestimate the actual error.

VI. RESULTS

For the intermediates and transition states depicted in the catalytic cycle in Fig. 1, we carried out density functional calculations that delivered the data presented in Fig. 2 and wave function calculations with various active orbital spaces for the analyses discussed below. For the sake of brevity we refer the reader to the Supplemental Material [28] for all technical details on these standard quantum chemical calculations.

For each of these key intermediates and transition states we then evaluate the cost of performing quantum phase estimation to chemical accuracy, both for small active spaces of 52–65 orbitals and larger ones, and then discuss runtimes and qubit requirements.

A. Selection of active orbitals

As mentioned in Sec. IV, the electronic Hamiltonian in Eq. (2) is parametrized by one- and two-electron integrals h_{ij} and h_{ijkl}, respectively, over the molecular orbitals of a restricted orbital subspace, the active space. For the selection of the active space (cf. Fig. 3), all strongly correlated orbitals must be identified (see the detailed discussion for our carbon dioxide fixation process in the supporting information). For this we relied on orbital entanglement measures which we applied in an automated procedure [47,49], but we emphasize that the importance of this choice for selecting the different orbital space sizes for this work is not crucial at all. Analysis of the resulting states in terms of these orbital entanglement measures, pair-orbital mutual information, and occupation numbers shows that the electronic ground states of the catalyst structures selected for our study are not dominated by static correlation. For such cases, the traditional quantum chemistry toolbox offers efficient and very accurate coupled cluster methods at the basis set limit, which represent a true challenge for quantum algorithms to compete with. Still, we
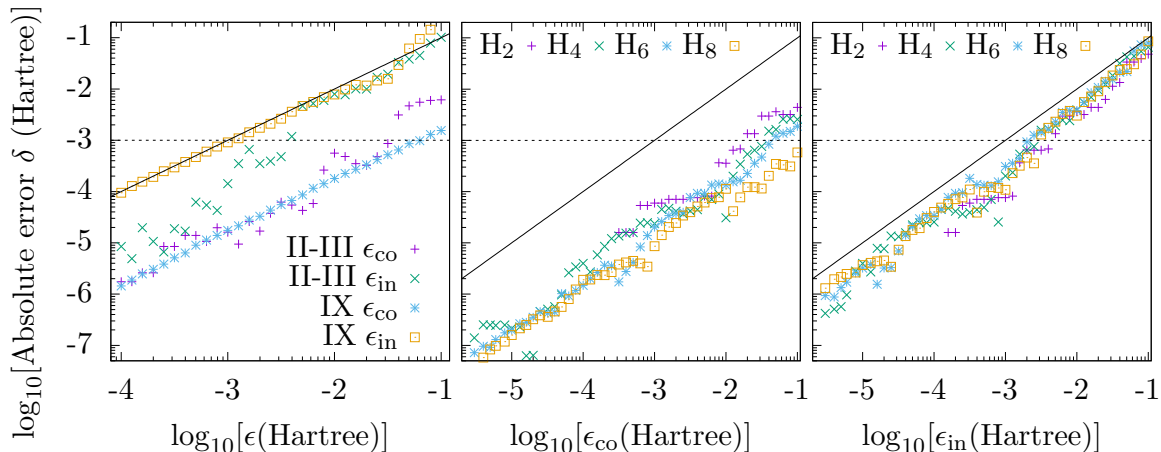
FIG. 4. Empirical absolute error δ (in Hartree) of the ground-state electronic energy resulting from applying the two truncation schemes to the two-electron integrals of the Hamiltonian at various thresholds ϵ_{co}, ϵ_{in}, evaluated from DMRG-CI calculations. The lines $\delta = \epsilon$ and $\delta = 1$mHartree are meant to guide the eye. (Left) Absolute error of the ground-state electronic DMRG-CI energy for the two truncation schemes for complex II–III with 6 active orbitals and complex IX with 16 active orbitals. (Middle) Absolute error in the ground-state electronic DMRG-CI energy of linear hydrogen chains of length 2, 4, 6, and 8 at different truncation thresholds ϵ_{co} for the coherent truncation scheme and (right) ϵ_{in} for the incoherent truncation.

seek to analyze a general approach in quantum computation that can deal with strong (static) correlation (and possibly with any correlation problem in the not too distant future). It is therefore not decisive for our resource analysis that the active space sizes chosen (see supporting information) neither obey nor exploit typical patterns of electron correlation. It was, however, still possible to select a small active space of strongly correlated orbitals on the size of 5 to 16 orbitals. The main purpose of this work is understanding the performance of quantum algorithms on an actual chemical problem of varying size. The varying size is the growing size of the orbital space in which the exact wave function is constructed. In order to construct a proper test bed, we are therefore forced to choose active space sizes that are arbitrary with respect to the proper balance of static and dynamic correlation. In other words, to grow the active space sizes beyond the classical limit, we had to add weakly correlated orbitals. Naturally it then became increasingly difficult to select the active spaces based on orbital entanglement alone (because all such orbitals show similarly weak entanglement entropy measures). In this study we therefore resorted to criteria which ensure a reproducible selection that is, at the same time, reasonable from a chemical point of view. We chose three different active space sizes (small, intermediate, and large) for each molecular structure of the catalytic cycle. The first active space comprises those orbitals selected based on orbital entanglement criteria. Note that this small active space is identical to the one employed in the CASSCF calculations that produced the molecular orbitals from which the Hamiltonian parameters h_{ij} and h_{ijkl} were calculated. The intermediate active space includes all the orbitals of the small active space and then in addition the valence orbitals of ruthenium and the ligands (excluding the triphos ligand and, if present, solvent molecules), as well as the orbitals involved in ligand-metal bonding. For the largest active space we further supplemented these orbitals with the π and π^* orbitals on the triphos ligand, as they form a

well-defined and easily identifiable set of orbitals. Apart from these three active spaces per catalyst structure, we created even larger active spaces for structure XVIII, for which we additionally selected three active spaces of size $n = 100$, 150, and 250 spatial molecular orbitals. Since a manual selection is pointless for such large active spaces of mostly dynamically correlated orbitals, we selected the $n/2$ occupied and $n/2$ virtual orbitals around the Fermi level. For each of these active spaces, we obtained the one- and two-electron integrals h_{ij} and h_{ijkl} via a four-index transformation (see also Fig. 3). These integrals then served as input to the quantum algorithms.

B. Resource estimates

We now evaluate the cost of phase estimation to chemical accuracy for various structures in the carbon capture catalytic cycle of Sec. II. The active spaces of the molecules considered in this section are in the range of 52–65 orbitals, although we tabulate more results for 2–250 orbitals in the Supplemental Material. In addition, we also evaluate the scaling of cost with respect to the active space size N while keeping the number of atoms fixed, and find a different scaling law of $M \sim N^{2.5}$ for the number of eigenvectors compared to that of $M \sim N \log N$ [55] when increasing the number of atoms. We find, as shown in Table I, that the gate cost for systems in the 52-orbital to 65-orbital range is on the order of 10^{10} Toffoli gates, using about 4000 qubits.

The cost depends on the truncation threshold ϵ_{in} as plotted in Fig. 5. At the highly aggressive threshold of $\epsilon_{in} = 100$ mHartree, the rank of these examples roughly matches the parameters fit by Peng *et al.* [55] to three-dimensional hydrocarbons for 54 orbitals. However, the numerically computed shift in energy at that threshold can exceed chemical accuracy, following Fig. 4, and should be interpreted as a most optimistic cost estimate. Therefore, we choose $\epsilon_{in} = 1$ mHartree,

TABLE I. Number of Toffoli gates for estimating an energy level to an error of 1 mHartree using a truncation threshold of $\epsilon_{in} = 1$ mHartree for the largest active spaces of structures in the catalytic cycle considered here. Our approach allows for a trade-off between the number of logical qubits required and the Toffoli count.

Structure	Orbitals	Electrons	R	M	α_{DF} per Hartree	Using fewer qubits		Using fewer Toffolis	
						Qubits	Toffolis/10^{10}	Qubits	Toffolis/10^{10}
I	52	48	613	23 566	177.3	3400	1.3	6900	1.1
II	62	70	734	33 629	374.4	4200	3.6	8400	3.1
II–III	65	74	783	38 122	416.0	4400	4.5	8900	3.7
V	60	68	670	29 319	371.1	4100	3.3	8200	2.9
VIII	65	76	794	39 088	425.7	4400	4.6	8900	3.8
VIII–IX	59	72	666	29 286	384.4	4000	3.4	8000	2.9
IX	62	68	638	28 945	396.6	4200	3.5	8400	3.1
XVIII	56	64	705	29 594	293.5	3700	2.5	7400	2.1

which our numerical simulations indicate closely reflect chemical accuracy. Irrespective of the truncation scheme, we note that the logarithmic scaling with $1/\epsilon_{in}$ means that the Toffoli costs vary by at most factor of 5 between these extremes. By varying the active space size between 2 and 250 orbitals for the different catalyst structures, we find in Fig. 6 that the Toffoli cost of phase estimation scales with $\sim N^{3.25}$. The exponent is a combination of two factors: Across all configurations of the catalytic cycle, the number of eigenvectors scales with $M \sim N^{2.5}$, hence $c_W \sim \sqrt{MN} \sim N^{1.75}$, and the normalizing constant $\alpha_{DF} \sim N^{1.5}$ Hartree.

We next demonstrate the extent of our algorithmic improvements by applying our techniques to the 54-orbital representation of the FeMoco active site of nitrogenase that we considered previously [11], and comparing costs with prior art based on Trotterization and qubitization of the single-factorized representation. As seen in Table II, our Toffoli cost following Eq. (19) is on the order of 1.22×10^{10} using 3600 qubits. Assuming that the number of logical qubits is not a limiting factor, this is a dramatic improvement over the 6.0×10^{14} T gates and 142 qubits in our original estimate [11], and a significant improvement over the 1.2×10^{12} Toffoli gates of the single-factorized approach [77], where $\alpha_{CD} = 3.6 \times 10^4$ Hartree with 3.0×10^5 unique nonzero terms, and a rank of

$R \sim 200$ was claimed to be sufficient to achieve chemical accuracy with respect to a CCSD ansatz. This is seen in Table III where, for the sake of comparison, we choose a truncation threshold of $\epsilon_{in} = 73$ mHartree to normalize the rank between our results and the single-factorized approach. Moreover, our double-factorized approach also improves upon the 2.3×10^{11} Toffoli gates of highly optimized implementations [77] based on the unfactorized Hamiltonian Eq. (2), which has $\alpha = 9.9 \times 10^3$ Hartree with 4.4×10^5 unique nonzero terms. Our improvement largely stems from a normalizing factor α_{DF} that is 33 to 120 times smaller as the number of terms in all approaches at this threshold are roughly equal. This trend also applies to the various carbon fixation catalyst structures that we consider when we perform a similar comparison but instead use the same incoherent truncation scheme for all examples at a more conservative error threshold of $\epsilon_{in} = 1$ mHartree.

C. Runtimes and qubit counts

We finally relate these gate count estimates to expected runtimes on future quantum computers. Exact runtime will, of course, depend on details of the system and error correction schemes. In our previous analysis [11], we optimistically
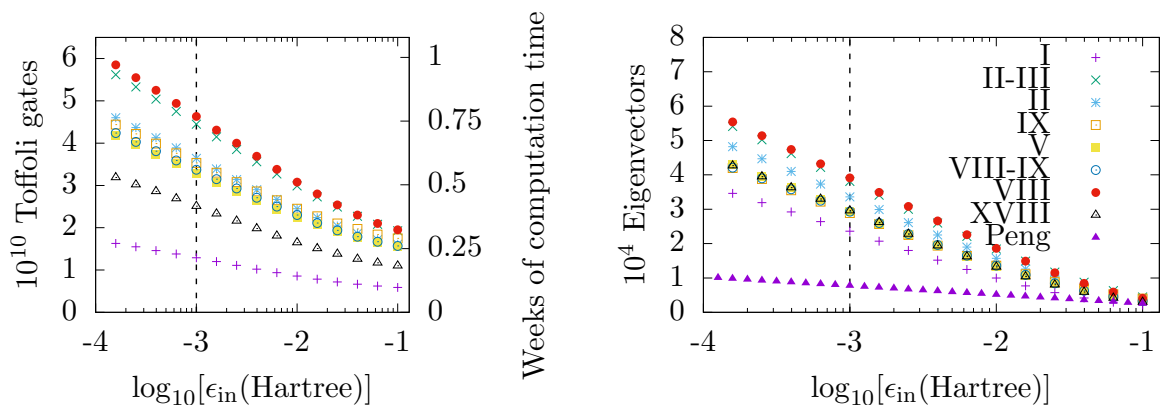

FIG. 5. (Left) Toffoli cost versus truncation threshold for phase estimation to a precision of 1 mHartree in the double-factorized representation of the catalytic cycle in Fig. 1 with a number of orbitals listed in Table I. Computation time is assumed to be 10 μs between each Toffoli, or 100 kHz. (Right) Number of eigenvector in the double-factorized representation. The data points by Peng *et al.* are for three-dimensional hydrocarbons [55] with 54 orbitals.
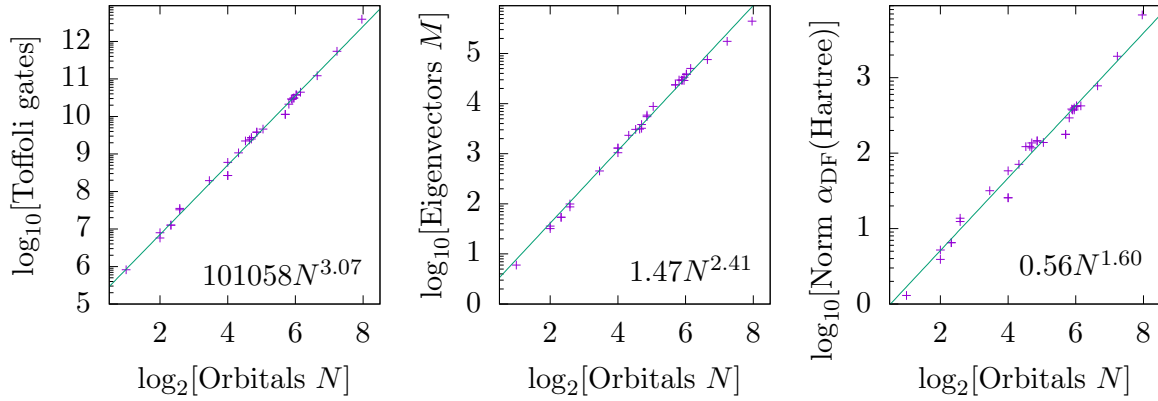
FIG. 6. (Left) Scaling of minimized Toffoli cost across all configurations of the catalytic cycle with respect to active space size, for performing phase estimation to a precision of 1 mHartree in the double-factorized representation. The fitted power-law curve has an exponent arising from the product $\alpha_{DF}c_W \sim \alpha_{DF}\sqrt{MN}/\Delta_E$ in Eqs. (1) and (19) (middle) where M is the number of eigenvectors and (right) the normalizing constant is α_{DF}.

assumed gate times of 100 ns for fault tolerant logical gate operations, which may be a long-term achievable goal. As a more realistic assumption for mid-term fault tolerant quantum computers we now expect that the physical gate times in current quantum computer architectures range from tens of nanoseconds for solid state qubits to tens of microseconds for ion traps. Realizing fault tolerance by using quantum error correction with the surface code [78], will lead to logical gate times for a Toffoli gate of about 10 μs to 10 ms depending on the architecture. The lower estimate of 10 μs means that 10^{10} Toffoli gates correspond to a runtime of 28 h or about a day, while the upper estimate of 10 ms would correspond to several years. These considerations show that fast gate times will be essential for realistic quantum computation for chemical catalysis.

The above estimates do not explicitly consider the cost of layout, i.e., the mapping onto a nearest neighbor planar square lattice topology of error corrected logical qubits. We argue that this overhead is negligible, since the subroutine with dominant cost, the table lookup discussed in the Supplemental Material is based on FANOUT operators [79] and maps well to this topology. Moreover, we here assume that any overhead of layout as well as Clifford gates are included in the assumed gate time for a Toffoli gate, as these dominant FANOUT operations may be implemented in a constant

Clifford depth of 4 in parallel with the sequentially applied Toffolis.

Fault tolerant gates also have an overhead in the number of qubits, with hundreds to thousands of physical qubits needed per logical qubit [11,78], depending on the quality of the qubits. 4000 logical qubits will thus correspond to millions of physical qubits. This implies the need for a scalable quantum computer architecture, scaling to millions of qubits.

D. State preparation

To determine the ground state energy using phase estimation, it is required to prepare a trial state $|\psi_{trial}\rangle$ which has a high overlap with the true ground state $|\psi_0\rangle$ of the Hamiltonian H as discussed in Sec. IV. While the exact ground state is unknown, we used state-of-the-art DMRG calculations to obtain an approximate ground state $|\tilde{\psi}_0\rangle$ for each of our systems (see the Supplemental Material [28] for further details). An approximate configuration interaction wave function obtained by reconstruction [80] from the corresponding matrix product state wave function optimized with DMRG served to provide an overlap with the trial state prepared on the quantum computer. Since the trial state is chosen to be HF determinant for this case, the overlap is given by the square of the coefficient in front of the Hartree-Fock determinant in the configuration

TABLE II. Comparison of our new double-factorization approach for H_{DF} applied to the FeMoco active site of nitrogenase ($N = 54$) with prior approaches based on Trotterization [11] or qubitization [22] using the unfactorized H or single-factorized H_{CD} Hamiltonian, and also for the VIII structure in the catalytic cycle ($N = 65$) where all examples apply the incoherent truncation scheme with the same threshold of $\epsilon_{in} = 1$ mHartree.

Structure	Approach	α / Hartree	Terms	Qubits	Toffoli gates	Comments
FeMoco	Qubitization H_{DF}	300.5	1.3×10^6	3600	2.3×10^{10}	$\epsilon_{in} = 1$ mHartree
	Qubitization H_{DF}	296.9	2.8×10^5	3600	1.22×10^{10}	Optimistic $\epsilon_{in} = 73$ mHartree
	Trotterization H [11]	–	–	142	1.5×10^{14}	Optimistic Trotter number
	Qubitization H [22]	9.9×10^3	4.4×10^5	5100	2.3×10^{11}	Truncation evaluated by CCSD
	Qubitization H_{CD} [22]	3.6×10^4	4.0×10^5	3000	1.2×10^{12}	Truncation evaluated by CCSD
VIII	Qubitization H_{DF}	425.7	2.5×10^6	4600	4.6×10^{10}	$\epsilon_{in} = 1$ mHartree
	Qubitization H	1.1×10^4	2.2×10^6	11000	9.3×10^{11}	$\epsilon_{in} = 1$ mHartree
	Qubitization H_{CD}	4.2×10^4	1.3×10^6	5800	2.1×10^{12}	$\epsilon_{in} = 1$ mHartree

TABLE III. Scaling of cost in our double-factorization approach with truncation threshold for the FeMoco active site of nitrogenase ($N = 54$). For comparison, the last line has $R = 200$ which matches that used by Berry *et al.* [22].

ϵ_{in}/mHartree	Rank R	Eigenvectors M	Terms $M \times N$	α_{DF}/Hartree	Qubits	No. Toffoli gates
1	567	2.4×10^4	1.30×10^6	300.5	3600	2.3×10^{10}
10	371	1.33×10^4	7.2×10^5	300.0	3600	1.67×10^{10}
100	178	4.2×10^3	2.3×10^5	295.8	3600	1.16×10^{10}
73	200	5.2×10^3	2.8×10^5	296.9	3600	1.22×10^{10}

interaction expansion, i.e., the one with the first $N/2$ orbitals occupied where N is the number of electrons. We found that there is a large overlap for all of our systems, see Table IV. We expect this overlap to not shrink substantially with increasing precision of the approximate ground state obtained by DMRG and hence preparing the dominant single-determinant state is sufficient for the molecular structures considered in this work. Recall that the Ru complexes selected for our study do not exhibit strong multireference character. If this overlap had turned out to be small, we would defer to Ref. [81] for options on how to prepare a multideterminant initial state in order to boost the success probability of phase estimation.

VII. CONCLUSIONS

In this work we considered computational catalysis leveraged by quantum computing. We rely on accurate error bounds on the electronic energy accessible in quantum algorithms to obtain sufficiently accurate data for intermediate and transition state structures of a catalytic cycle. In particular, (i) we considered decisive steps of a synthetic catalyst that is known to convert carbon dioxide to methanol (and for which traditional work based on density functional theory had already been reported alongside the experimental results [27]), (ii) we presented a quantum algorithm in the qubitization framework that exploits a double-factorized electronic structure representation, which significantly reduces the runtime, (iii) we validated the truncation schemes by comparison with density matrix renormalization group calculations, (iv) we confirmed that starting the quantum algorithm from a single determinant initial state has high success probability for the carbon dioxide functionalization process studied in this work, and (v) we calculated realistic resource estimates for mid-term

scalable quantum computers and related them to expected runtimes.

A pressing challenge in the cost of quantum simulation for large molecular systems is the rapid growth in the number of four-index two-electron integrals. By using integral decomposition techniques in the double-factorized representation [56,82], we simultaneously minimize two key parameters governing the cost of qubitization: the number of coefficients that must be loaded into the quantum computer and a certain bound α_{DF} on the spectral norm of the Hamiltonian. Compared to other simulation techniques such as Trotterization, qubitization also enables a large trade-off in non-Clifford gate count by using additional qubits. In the case of estimating an energy level to chemical accuracy, we found that our approach has a Toffoli gate cost with respect to active space size N, while keeping the number of atoms fixed, that scales like $\sim N^{3.25}$. When the number of atoms also grows with N, our technique promises a Toffoli cost scaling like $\sim \alpha_{\mathrm{DF}} N \log(N)$, assuming the empirically fitted sparsity of double factorization by other authors [55], and where $\alpha_{\mathrm{DF}} \sim N^{1.5}$ based on hydrogen chain benchmarks. When applied to structures in the catalytic cycle with 52–65 orbitals, our approach requires roughly 10^{10}–10^{11} Toffoli gates and ~ 4000 logical qubits. Note that the benefit of our reduced Toffoli count can far outweigh the cost of using more logical qubits—the physical qubit count of the overall algorithm is typically dominated by that required for the fault-tolerant distillation of non-Clifford gates. Under realistic assumptions on the performance of mid-term quantum hardware, this corresponds to a runtime of a few weeks for a single calculation. While this seems similar to the results of our previous paper [11], we here use much more conservative and realistic estimates for the clock speeds of mid-term quantum computers, while Ref. [11] assumed ambitious specifications for the long term.

TABLE IV. Overlap $|\langle \tilde{\psi}_0 | \psi_{\mathrm{trial}} \rangle|^2$ of the dominant single-determinant state $|\psi_{\mathrm{trial}} \rangle$ with the approximate ground state $|\tilde{\psi}_0 \rangle$ obtained with DMRG for the complexes considered in this work. The type of molecular orbital (MO) basis is given as well as the size of active space of the DMRG calculation in terms of orbitals and electrons.

Complex	MO basis	Orbitals	Electrons	$|\langle \tilde{\psi}_0 | \psi_{\mathrm{trial}} \rangle|^2$
I	CAS(4,5)SCF	52	48	0.847
II	CAS(8,6)SCF	62	70	0.848
II–III	CAS(8,6)SCF	65	74	0.848
V	CAS(12,11)SCF	60	68	0.841
VIII	CAS(2,2)SCF	65	76	0.869
VIII–IX	CAS(4,4)SCF	59	72	0.863
IX	CAS(16,16)SCF	62	68	0.807
XVIII	CAS(4,4)SCF	56	64	0.889

From our results it is evident that a future universal quantum computer can only lead the range of electronic structure methods if further developments provide faster algorithms that can treat much larger active spaces in order to be competitive. We emphasize that, although no logical qubit has been realized experimentally so far, a universal quantum computer that can represent a state on a few thousand logical qubits would revolutionize electronic structure theory as it bears the potential to provide an accurate exact-diagonalization energy for a moderately sized molecular system of on the order of 100 atoms in the full single-particle basis set. As a consequence, any necessarily approximate *a priori* or *a posteriori* correction for the nagging dynamic correlation problem (that arises solely from the choice of a reduced-dimensional active space) would no longer be needed. It is clear that such a calculation would be an ultimate goal, but also that it would present new technical challenges.

Going to an order of magnitude larger active space sizes as considered here would increase the quantum computational requirements by three orders of magnitude. Hence, despite the advances reported in this work, it is obvious that further algorithmic improvements of quantum algorithms are urgently needed in order to get timings down so that a quantum computer can become superior to traditional approaches, i.e., to eventually demonstrate a quantum advantage in real-world chemistry applications. Given the innovation rate in quantum algorithms for chemistry of many orders of magnitude over the past years, we are confident for this to happen. One direction is exploring novel sparse representations. Another promising avenue is exploiting the fact that the number of electrons may grow much smaller than the size of the single-particle basis set when expanding the latter to incorporate dynamical correlations. That can lead to significant savings compared to the algorithms presented here.

The integral files used in this work are hosted on Zenodo [83].

[1] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).

[2] K. M. Svore and M. Troyer, The quantum future of computation, IEEE Comput. **49**, 21 (2016).

[3] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, Quantum chemistry in the age of quantum computing, Chem. Rev. **119**, 10856 (2019).

[4] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, Rev. Mod. Phys. **92**, 015003 (2020).

[5] D. S. Abrams and S. Lloyd, Simulation of Many-Body Fermi Systems on a Universal Quantum Computer, Phys. Rev. Lett. **79**, 2586 (1997).

[6] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Simulated quantum computation of molecular energies, Science **309**, 1704 (2005).

[7] L. Veis and J. Pittner, Quantum computing applied to calculations of molecular energies: CH2 benchmark, J. Chem. Phys. **133**, 194106 (2010).

[8] L. Veis, J. Višňák, T. Fleig, S. Knecht, T. Saue, L. Visscher, and J. Pittner, Relativistic quantum chemistry on quantum computers, Phys. Rev. A **85**, 030304(R) (2012).

[9] L. Veis and J. Pittner, Quantum computing approach to non-relativistic and relativistic molecular energy calculations, in *Quantum Information and Computation for Chemistry* (John Wiley and Sons, New York, 2014), pp. 107.

[10] K. Bourzac, Chemistry is quantum computing's killer app, C&EN **95**, 27 (2017).

[11] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, Elucidating reaction mechanisms on quantum computers, Proc. Natl. Acad. Sci. U.S.A. **114**, 7555 (2017).

[12] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, Phys. Rev. Lett. **69**, 2863 (1992).

[13] A. Baiardi and M. Reiher, The density matrix renormalization group in chemistry and molecular physics: recent developments and new challenges, J. Chem. Phys. **152**, 040903 (2020).

[14] G. H. Booth, A. J. W. Thom, and A. Alavi, Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space, J. Chem. Phys. **131**, 054106 (2009).

[15] G. H. Low and I. L. Chuang, Hamiltonian simulation by qubitization, Quantum **3**, 163 (2019).

[16] G. H. Low and I. L. Chuang, Optimal Hamiltonian Simulation by Quantum Signal Processing, Phys. Rev. Lett. **118**, 010501 (2017).

[17] J. Haah, M. Hastings, R. Kothari, and G. H. Low, Quantum algorithm for simulating real time evolution of lattice Hamiltonians, in *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, FOCS '18 (IEEE, Washington, DC, 2018), pp. 350.

[18] G. H. Low and N. Wiebe, Hamiltonian simulation in the interaction picture, arXiv:1805.00675.

[19] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, A Theory of Trotter Error, Phys. Rev. X **11**, 011020 (2021).

[20] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven, Encoding Electronic Spectra in Quantum Circuits with Linear T Complexity, Phys. Rev. X **8**, 041015 (2018).

[21] G. H. Low, V. Kliuchnikov, and L. Schaeffer, Trading T-gates for dirty qubits in state preparation and unitary synthesis, arXiv:1812.00954.

[22] D. W. Berry, C. Gidney, M. Motta, J. R. McClean, and R. Babbush, Qubitization of arbitrary basis quantum chemistry

leveraging sparsity and low rank factorization, Quantum **3**, 208 (2019).

[23] C. Hepburn, E. Adlen, J. Beddington, E. A. Carter, S. Fuss, N. M. Dowell, J. C. Minx, P. Smith, and C. K. Williams, The technological and economic prospects for CO_2 utilization and removal, Nature (London) **575**, 87 (2019).

[24] Y. Wu, Z. Jiang, X. Lu, Y. Liang, and H. Wang, Domino electroreduction of CO_2 to methanol on a molecular catalyst, Nature (London) **575**, 639 (2019).

[25] F. Li, A. Thevenon, A. Rosas-Hernández, Z. Wang, Y. Li, C. M. Gabardo, A. Ozden, C. T. Dinh, J. Li, Y. Wang, J. P. Edwards, Y. Xu, C. McCallum, L. Tao, Z.-Q. Liang, M. Luo, X. Wang, H. Li, C. P. O'Brien, C.-S. Tan, D.-H. Nam, R. Quintero-Bermudez, T.-T. Zhuang, Y. C. Li, Z. Han, R. D. Britt, D. Sinton, T. Agapie, J. C. Peters, and E. H. Sargent, Molecular tuning of CO_2-to-ethylene conversion, Nature (London) **577**, 509 (2020).

[26] S. Kar, J. Kothandaraman, A. Goeppert, and G. K. S. Prakash, Advances in catalytic homogeneous hydrogenation of carbon dioxide to methanol, J. CO_2 Util. **23**, 212 (2018).

[27] S. Wesselbaum, V. Moha, M. Meuresch, S. Brosinski, K. M. Thenert, J. Kothe, T. vom Stein, U. Englert, M. Hölscher, J. Klankermayer, and W. Leitner, Hydrogenation of carbon dioxide to methanol using a homogeneous ruthenium–Triphos catalyst: From mechanistic investigations to multiphase catalysis, Chem. Sci. **6**, 693 (2015).

[28] See Supplemental Material at http://link.aps.org/supplemental/ 10.1103/PhysRevResearch.3.033055 for details on the electronic structure calculations and the qubitization algorithm presented in this work as well as tabulated resource estimates for active spaces up to 250 orbitals.

[29] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Challenges for density functional theory, Chem. Rev. **112**, 289 (2012).

[30] N. Mardirossian and M. Head-Gordon, Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals, Mol. Phys. **115**, 2315 (2017).

[31] T. Weymuth, E. P. A. Couzijn, P. Chen, and M. Reiher, New benchmark set of transition-metal coordination reactions for the assessment of density functionals, J. Chem. Theory Comput. **10**, 3092 (2014).

[32] G. N. Simm and M. Reiher, Systematic error estimation for chemical reaction energies, J. Chem. Theory Comput. **12**, 2762 (2016).

[33] T. Husch, L. Freitag, and M. Reiher, Calculation of ligand dissociation energies in large transition-metal complexes, J. Chem. Theory Comput. **14**, 2456 (2018).

[34] J. Proppe, T. Husch, G. N. Simm, and M. Reiher, Uncertainty quantification for quantum chemical models of complex reaction networks, Faraday Discuss. **195**, 497 (2016).

[35] Y. Zhao and D. G. Truhlar, A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions, J. Chem. Phys. **125**, 194101 (2006).

[36] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys. Rev. Lett. **77**, 3865 (1996).

[37] J. P. Perdew, M. Ernzerhof, and K. Burke, Rationale for mixing exact exchange with density functional approximations, J. Chem. Phys. **105**, 9982 (1996).

[38] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, 2nd ed. (Wiley, Chichester, 2006).

[39] F. Jensen, *Introduction to Computational Chemistry* (Wiley, New York, 2007).

[40] G. N. Simm, A. C. Vaucher, and M. Reiher, Exploration of reaction pathways and chemical transformation networks, J. Phys. Chem. A **123**, 385 (2019).

[41] Q. Ma and H.-J. Werner, Explicitly correlated local coupled-cluster methods using pair natural orbitals, Wiley Interdiscip. Rev. Comput. Mol. Sci. **8**, e1371 (2018).

[42] I. Fdez. Galván, M. Vacher, A. Alavi, C. Angeli, F. Aquilante, J. Autschbach, J. J. Bao, S. I. Bokarev, N. A. Bogdanov, R. K. Carlson, L. F. Chibotaru, J. Creutzberg, N. Dattani, M. G. Delcey, S. S. Dong, A. Dreuw, L. Freitag, L. M. Frutos, L. Gagliardi, F. Gendron, A. Giussani, L. González, G. Grell, M. Guo, C. E. Hoyer, M. Johansson, S. Keller, S. Knecht, G. Kovačević, E. Källman, G. Li Manni, M. Lundberg, Y. Ma, S. Mai, J. P. Malhado, P. Å. Malmqvist, P. Marquetand, S. A. Mewes, J. Norell, M. Olivucci, M. Oppel, Q. M. Phung, K. Pierloot, F. Plasser, M. Reiher, A. M. Sand, I. Schapiro, P. Sharma, C. J. Stein, L. K. Sørensen, D. G. Truhlar, M. Ugandi, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, O. Weser, T. A. Wesołowski, P.-O. Widmark, S. Wouters, A. Zech, J. P. Zobel, and R. Lindh, OpenMolcas: From source code to insight, J. Chem. Theory Comput. **15**, 11 (2019).

[43] K. D. Vogiatzis, D. Ma, J. Olsen, L. Gagliardi, and W. A. de Jong, Pushing configuration-interaction to the limit: Towards massively parallel MCSCF calculations, J. Chem. Phys. **147**, 184111 (2017).

[44] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach, Chem. Phys. **48**, 157 (1980).

[45] K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert, Are atoms intrinsic to molecular electronic wavefunctions? I. The FORS model, Chem. Phys. **71**, 41 (1982).

[46] S. Keller, K. Boguslawski, T. Janowski, M. Reiher, and P. Pulay, Selection of active spaces for multiconfigurational wavefunctions, J. Chem. Phys. **142**, 244104 (2015).

[47] C. J. Stein and M. Reiher, Automated selection of active orbital spaces, J. Chem. Theory Comput. **12**, 1760 (2016).

[48] C. J. Stein, V. von Burg, and M. Reiher, The delicate balance of static and dynamic electron correlation, J. Chem. Theory Comput. **12**, 3764 (2016).

[49] C. J. Stein and M. Reiher, autoCAS: A program for fully automated multiconfigurational calculations, J. Comput. Chem. **40**, 2216 (2019).

[50] L. González and R. Lindh, *Quantum Chemistry and Dynamics of Excited States: Methods and Applications* (Wiley, New York, 2020).

[51] E. Fromager, J. Toulouse, and H. J. A. Jensen, On the universality of the long-/short-range separation in multiconfigurational density-functional theory, J. Chem. Phys. **126**, 074111 (2007).

[52] E. D. Hedegård, S. Knecht, J. S. Kielberg, H. J. Aagaard Jensen, and M. Reiher, Density matrix renormalization group with efficient dynamical electron correlation through range separation, J. Chem. Phys. **142**, 224108 (2015).

[53] E. D. Hedegård, J. Toulouse, and H. J. A. Jensen, Multiconfigurational short-range density-functional theory for open-shell systems, J. Chem. Phys. **148**, 214103 (2018).

[54] H. Luo and A. Alavi, Combining the transcorrelated method with full configuration interaction quantum Monte Carlo:

Application to the homogeneous electron gas, J. Chem. Theory Comput. **14**, 1403 (2018).

[55] B. Peng and K. Kowalski, Highly efficient and scalable compound decomposition of two-electron integral tensor and its application in coupled cluster calculations, J. Chem. Theory Comput. **13**, 4179 (2017).

[56] F. Aquilante, L. Boman, J. Boström, H. Koch, R. Lindh, A. S. de Merás, and T. B. Pedersen, Cholesky decomposition techniques in electronic structure theory, in *Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications*, Challenges and Advances in Computational Chemistry and Physics, edited by R. Zalesny, M. G. Papadopoulos, P. G. Mezey, and J. Leszczynski (Springer Netherlands, Dordrecht, 2011), pp. 301.

[57] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 1 (2014).

[58] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, Nat. Commun. **10**, 1 (2019).

[59] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[60] M. Kühn, S. Zanker, P. Deglmann, M. Marthaler, and H. Weiß, Accuracy and resource estimations for quantum chemistry on a near-term quantum computer, J. Chem. Theory Comput. **15**, 4764 (2019).

[61] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, Phys. Rev. A **92**, 042303 (2015).

[62] N. Wiebe and C. Granade, Efficient Bayesian Phase Estimation, Phys. Rev. Lett. **117**, 010503 (2016).

[63] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 1st ed. (Cambridge University Press, Cambridge, 2004).

[64] W. Górecki, R. Demkowicz-Dobrzański, H. M. Wiseman, and D. W. Berry, π-Corrected Heisenberg Limit, Phys. Rev. Lett. **124**, 030501 (2020).

[65] D. Wecker, M. B. Hastings, N. Wiebe, B. K. Clark, C. Nayak, and M. Troyer, Solving strongly correlated electron models on a quantum computer, Phys. Rev. A **92**, 062318 (2015).

[66] M. Suzuki, Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations, Phys. Lett. A **146**, 319 (1990).

[67] G. H. Low, Hamiltonian simulation with nearly optimal dependence on spectral norm, in *Proceedings of the 51st Annual ACM Symposium on Theory of Computing - STOC '19* (ACM, New York, 2019), pp. 491.

[68] A. M. Childs and N. Wiebe, Hamiltonian simulation using linear combinations of unitary operations, Quantum Info. Comput. **12**, 901 (2012).

[69] M. Reiher and A. Wolf, *Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science*, 2nd ed. (Wiley-VCH, Weinheim, 2015).

[70] D. Poulin, A. Kitaev, D. S. Steiger, M. B. Hastings, and M. Troyer, Quantum Algorithm for Spectral Measurement with a Lower Gate Count, Phys. Rev. Lett. **121**, 010501 (2018).

[71] D. W. Berry, M. Kieferová, A. Scherer, Y. R. Sanders, G. H. Low, N. Wiebe, C. Gidney, and R. Babbush, Improved techniques for preparing eigenstates of fermionic Hamiltonians, npj Quantum Inf **4**, 22 (2018).

[72] D. Maslov, Optimal and asymptotically optimal NCT reversible circuits by the gate types, Quantum Inf. Comput. **16**, 1096 (2016).

[73] C. Gidney and A. G. Fowler, Efficient magic state factories with a catalyzed $|CCZ\rangle$ to $2|T\rangle$ transformation, Quantum **3**, 135 (2019).

[74] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doberty, and M. Troyer, The trotter step size required for accurate quantum simulation of quantum chemistry, Quantum Info. Comput. **15**, 361 (2015).

[75] M. Motta, E. Ye, J. R. McClean, Z. Li, A. J. Minnich, R. Babbush, and G. K.-L. Chan, Low rank representations for quantum simulation of electronic structure, npj Quantum Inf. **7**, 83 (2021).

[76] M. Motta, J. Shee, S. Zhang, and G. K.-L. Chan, Efficient ab initio auxiliary-field quantum Monte Carlo calculations in Gaussian bases via low-rank tensor decomposition, J. Chem. Theory Comput. **15**, 3510 (2019), pMID: 31091103.

[77] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Simulating Hamiltonian Dynamics with a Truncated Taylor Series, Phys. Rev. Lett. **114**, 090502 (2015).

[78] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, Phys. Rev. A **86**, 032324 (2012).

[79] P. Pham and K. M. Svore, A 2D nearest-neighbor quantum architecture for factoring in polylogarithmic depth, Quantum Info. Comput. **13**, 937 (2013).

[80] K. Boguslawski, K. H. Marti, and M. Reiher, Construction of CASCI-type wave functions for very large active spaces, J. Chem. Phys. **134**, 224101 (2011).

[81] N. M. Tubman, C. Mejuto-Zaera, J. M. Epstein, D. Hait, D. S. Levine, W. Huggins, Z. Jiang, J. R. McClean, R. Babbush, M. Head-Gordon, and K. B. Whaley, Postponing the orthogonality catastrophe: efficient state preparation for electronic structure simulations on quantum devices, arXiv:1809.05523.

[82] N. H. F. Beebe and J. Linderberg, Simplifications in the generation and transformation of two-electron integrals in molecular calculations, Int. J. Quantum Chem. **12**, 683 (1977).

[83] V. von Burg, G. H. Low, T. Häner, D. S. Steiger, M. Reiher, M. Roetteler, and M. Troyer, dataset for "Quantum computing enhanced computational catalysis (molecular integrals)", 2021, doi: 10.5281/zenodo.4769113.