communications physics

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42005-025-02305-9

Accurate quantum-centric simulations of intermolecular interactions

Danil Kaliakin[1], Akhil Shajan[1,2], Fangchun Liang[1], Javier Robledo Moreno[3], Zhen Li[1], Abhishek Mitra[1], Mario Motta[3], Caleb Johnson[3], Abdullah Ash Saki[3], Susanta Das[1], Iskandar Sitdikov[3], Antonio Mezzacapo[3] & Kenneth M. Merz Jr[1,2]

Modelling and simulating non-covalent interactions is challenging, as they are inherently weak, dynamic, and system-specific. Common predictive methods often require trading the accuracy for reducing the otherwise cumbersome computational cost. To date, the most accurate approaches, achieving chemical accuracy, rely on quantum mechanical descriptions of non-covalent interactions, which limits their scalability. Whether quantum computing could overcome these limitations is still unclear, as such methods need to be redesigned for quantum hardware. Here, we take the first step in this direction by presenting quantum-centric simulations of non-covalent interactions using a supramolecular approach for binding energy calculations. We use a sample-based quantum diagonalization (SQD) approach to simulate the potential energy surfaces (PES) of the water and methane dimers, featuring hydrogen bond and dispersion interactions, respectively. We benchmark our quantum simulations (27- and 36-qubit circuits) against classical methods, registering deviations within 1.000 kcal/mol from the leading ones. Finally, we test the limits of the quantum methods for capturing dispersion interactions with an experiment on 54 qubits. Beyond reaching state-of-the-art accuracy, our work lays out a framework for electronic structure calculations of non-covalent interactions on quantum hardware.

The accurate treatment of non-covalent interactions[1,2] is important in the biological, chemical, and pharmaceutical sciences[3,4]. Specifically, non-covalent interactions between hydrophobic species and hydrogen-bonded pairs play critical roles in biological processes, such as protein folding[5–9], membrane assembly[10], and cell signaling[11]. Understanding of non-covalent interactions is critical for drug discovery[12–16]. The correct modeling of these interactions along with solvation plays a key role in understanding chemical and biological processes[17].

Traditionally, quantum mechanical[18,19] methods have been used to study these systems to a high level of accuracy—so-called chemical accuracy (±1 kcal/mol from experiment). However, these calculations are quite expensive and approaches to accelerate these calculations continue to be explored using classical hardware[20–29]. Using the results from these calculations, force fields have been fine-tuned for wide use in molecular simulation studies of chemical and biological processes[30–34]. More recently, machine learning[35] methods built using accurate quantum chemical calculations on thousands of systems have appeared to study largely covalent interactions, but can be extended to non-covalent interactions at a good level of accuracy, but at a reduced computational cost. However, these latter methods build models that can struggle to study diverse systems outside of the training set and can be subject to overfitting[36,37].

Quantum computing exploits superposition and entanglement in attempt to solve complex problems faster than classical methods. Approaches like quantum Gibbs sampling, including quantum Metropolis sampling[38,39], and variational methods[40,41] offer promising advantages. Quantum computing based studies of non-covalent interactions, to a high level of accuracy and speed, would improve our ability to understand complex processes like drug binding, but would also allow for the development of large synthetic datasets that could be used to build even better force fields and quantum machine learning models. However, to date, quantum hardware has struggled to address these problems. In this work we demonstrate that quantum-centric supercomputing (QCSC)[42] combined with the sample-based quantum diagonalization (SQD) approach[43] allows for the study of intermolecular interactions.

QCSC is a computational paradigm, in which a quantum computer operates in concert with classical high-performance computing (HPC) resources. Classical processing carried out before, during, and after quantum computations allows for the introduction of quantum subroutines in

[1]Center for Computational Life Sciences, Lerner Research Institute, The Cleveland Clinic, Cleveland, OH, USA. [2]Department of Chemistry, Michigan State University, East Lansing, MI, USA. [3]IBM Quantum, IBM T.J. Watson Research Center, Yorktown Heights, NY, USA. e-mail: kmerz1@gmail.com

the workflow of classical HPC algorithms, to extract and amplify signal from noisy quantum devices, and to leverage quantum processors to execute a limited number of large quantum circuits.

The QCSC architecture enables scaling of computational capabilities, as exemplified by methods that use classical diagonalization in subspaces determined by quantum samples such as SQD[43] and QSCI[44]. The SQD method is developed based on QSCI. The SQD method uses a quantum device to sample electronic configurations from a quantum circuit approximating the ground state of a molecular Hamiltonian, and uses classical distributed HPC resources to post-process quantum measurements against known symmetries to obtain recovered configurations[43], as well as to solve the Schrödinger equation in the subspace spanned by the recovered configurations.

Sampling of electronic configurations from a quantum circuit is achieved with approximate wavefunction ansatzes. The most commonly used ansatz for eigenstates of electronic Hamiltonians is the unitary coupled cluster with single and double excitations (UCCSD) and it's approximate forms[45,46]. The exponential ansatz provides a compact, approximate description of infinite-order electronic excitations, critical for the correct physical descriptions of many-body screening and polarization effects[47]. The unitary coupled cluster method is more efficient than traditional coupled cluster models in the description of static correlation effects beyond the two-electron problems[46]. The UCCSD ansatz offers a path towards quantum advantage, because the cost of evaluating eigenstates of the UCCSD ansatz is polynomial with the system size on a quantum computer and factorial on a conventional computer[46]. However, the exact simulations of the UCCSD ansatz would require the circuits of high depth, which are not practical on NISQ devices.

The local unitary coupled cluster (LUCJ) ansatz offers the efficient approximation of UCCSD ansatz while significantly reducing the circuit depth[46]. In the present work the LUCJ ansatz allows us to get an approximation of the UCCSD electron configuration distribution within ~85 s for the methane dimer (16e,16o) and in ~ 229 s for the methane dimer (16e,24o) using the ibm_cleveland and ibm_kyiv devices, respectively. We also observe similar timings for systems with strong electron correlation such as [2Fe-2S] and [4Fe-4S] using IBM Heron quantum processing units. At the present we need to perform a self-consistent configuration recovery (S-CORE) procedure, which substantially increases the computational time, to recover the configurations that were corrupted by the noise introduced by the quantum computer. However, we believe with continuous improvement of the LUCJ ansatz parameter optimization procedure and further reduction of noise in future quantum devices will allow us to identify the dominant electron configurations of a given molecular systems with greater efficiency than fully classical approaches[43].

The SQD method recently allowed us to address instances of the electronic structure problem with up to 36 spatial orbitals using up to 77 qubits[43]. The QCSC workflows produced improvements over simulations using quantum computers in isolation—which have in the last decade, used up to a handful of qubits with limited accuracies[48–53]. The QCSC paradigm coupled with SQD enables the study of problems heretofore out of reach of quantum computers including static correlation in iron-sulfur complexes[43] as well as dynamical correlation as exemplified in the intermolecular interactions studied herein. Moreover, using the example of the [2Fe-2S] cluster, the previous study demonstrated the possibility of quantum circuits that can sample electronic configurations more efficiently than some classes of classical heuristics, yielding equal or better energies than classical heuristics at fixed subspace dimensions[43], showcasing the potential feasibility of future quantum advantage within the QCSC paradigm. The quantum advantage within the QCSC paradigm should be especially evident in larger active spaces and when calculations are performed further from equilibrium[43].

Past studies have reported the simulation of non-covalent interactions[54,55] using symmetry-adapted perturbation theory (SAPT). This method expresses the interaction energy through a perturbative treatment of the intermolecular potential[56–58], and requires the simulation of electronic structure of individual monomers on a quantum computer. In addition Anderson et al. demonstrated the possibility of simulations of coarse-grained intermolecular interactions on a quantum computer as well[59]. However, to date, predicting binding energies between monomers using the supermolecular approach, where the electronic structure of the entire dimer needs to be simulated on quantum hardware, has been an elusive target for quantum simulations, due to lack of accuracy and scale of conventional quantum approaches.

Herein, we present the supermolecular quantum-centric simulations of binding energies in non-covalently bound pairs of molecules featuring hydrogen bond and dispersion interactions. We simulate the potential energy surfaces (PES) of the water dimer and the methane dimer. Our water dimer simulations use 27-qubit circuits, while the methane dimer simulations use 36- and 54-qubit circuits. To assess the accuracy of our quantum solutions, we compare them against heat-bath configuration interaction (HCI)[60–63] in the case of (16e,24o) calculations, complete active space configuration interaction for the (16e,12o) and (16e,16o) calculations, as well as coupled-cluster singles, doubles and perturbative triples (CCSD(T))[64] performed for all of the studied instances. The latter is widely recognized as the gold standard for computing intermolecular interactions[65] to chemical accuracy. For the 27-qubit water dimer and the 36-qubit methane dimer simulations, we demonstrate that SQD energies agree with CASCI nearly exactly, while deviating from CCSD(T) within 1 kcal/mol in the equilibrium region of the PES. For the 54-qubit simulations of the methane dimer, we observe how the accuracy of the quantum solution can be systematically improved by increasing the number of sampled configurations.

We note that this study was dedicated to establishing whether the accurate simulation of noncovalent interactions is possible within the quantum-centric SQD workflow, while elucidation of the possibility of quantum advantage in simulations of non-covalent interactions would require series of subsequent studies benchmarking extensive number of various molecular systems identifying specific cases where such advantage is possible. We anticipate that the findings of the present paper will pave the way for these more extensive benchmarking studies. In these future studies we anticipate moving past just demonstration of SQD simulations of non-covalent interactions and address the possibility of quantum advantage.

In addition to showing the application of SQD in simulations of non-covalent interactions, the present work also achieved other important technical milestones for the SQD method. While the previous SQD study[43] demonstrated that the SQD total energy extrapolation using Hamiltonian variance can produce an accurate estimate of the total energy at the equilibrium geometry, the present paper shows that this approach can work efficiently in predicting the interaction all along the PES as well. The present work also demonstrated the diagonalization of the largest subspace with SQD to date. While the previous SQD study[43] demonstrated SQD calculation using a subspace of 1.00×10^8 configurations for the [4Fe-4S] cluster, the present study employs a subspace of 2.49×10^8 configurations in the case of (16e,24o) methane dimer system. Finally, the present study demonstrated how AVAS[66] can be successfully integrated with the SQD software stack along with the Qiskit addon: SQD[67] and PySCF[68–70], where AVAS enables efficient and convenient active space selection for SQD calculations. The integration of AVAS, which was initially demonstrated in the present work, was also crucial for our DMET SQD[71] and SQD IEF-PCM studies[72].

Results and Discussion

Water Dimer Simulations

Figure 1 shows the binding energy of the water dimer as a function of the oxygen-oxygen distance using SQD and CASCI. The SQD and CASCI potential energy surfaces closely align, deviating from each by less than 0.020 kcal/mol. This close alignment is an indication that both methods have accurately solved the Schrodinger equation in the active space. The active-space SQD and CASCI calculations cannot capture dynamical correlation from inactive orbitals. To quantify the extent of the active-space

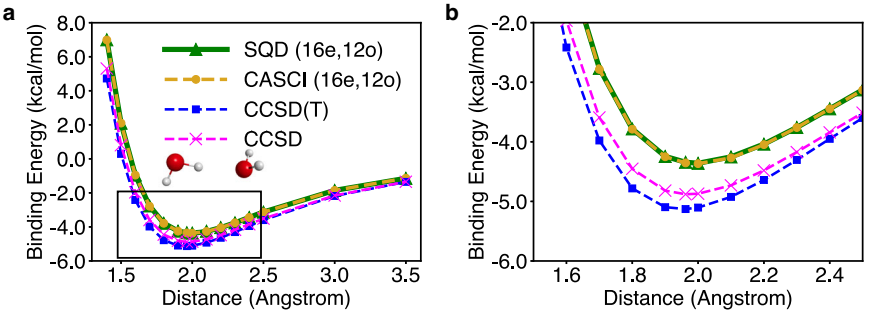

Fig. 1 | Binding energies of the water dimer along the PES, where distances between oxygen atoms range between 1.400 and 3.500 Å. a The entire range of bondlengths, and **b** a magnified region near equilibrium, highlighted in (**a**) as a black box. Light brown, blue, and magenta dashed lines with circle, square, and cross markers depict the PES calculated with the CASCI (16e,12o), CCSD(T), and CCSD methods, respectively. The solid green line with triangular markers depicts the PES calculated with the SQD (16e,12o).

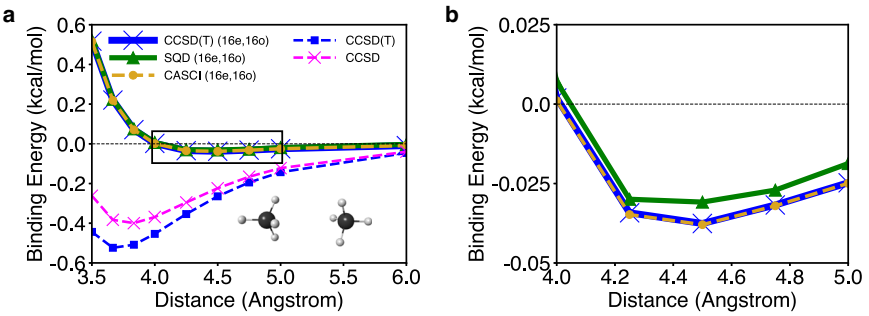

Fig. 2 | Binding energies of the methane dimer along the PES, where the distances between the carbon atoms range between 3.500 and 6.000 Å. Active-space simulations use (16e,16o). **a** The entire range of bondlengths, and **b** a magnified region near equilibrium, highlighted in (**a**) as a black box. Light brown, blue, and magenta dashed lines with circle, square, and cross markers depict the PES calculated with the CASCI (16e,16o), CCSD(T), and CCSD methods, respectively. The solid green line with triangular markers depicts the PES calculated with the SQD (16e,16o). The solid blue line represents CCSD(T) (16e,16o) calculations. The black horizontal dashed line indicates the zero value of the binding energy.

approximation, we also compute the potential energy surface using CCSD and CCSD(T) in the full aug-cc-pVQZ basis. The perturbative triples do not have a drastic effect on the binding energy between water monomers and the close agreement between CCSD and CCSD(T) calculations is shown in Fig. 1b. The excellent agreement between CCSD and CCSD(T) in the full basis and between SQD and CASCI in the active space indicates that the differences between SQD and CCSD(T) are due to the active-space approximation underlying the former. The CCSD(T) and SQD potential energy curves are in reasonable agreement with each other, the highest deviation being observed at 1.400 Å and corresponding to 2.263 kcal/mol. Despite this reasonable agreement and the ability of SQD to capture hydrogen bonding, there are quantitative differences in the predicted binding energies, -5.129 kcal/mol and -4.366 kcal/mol for CCSD(T) and SQD respectively, and the lowest-energy distances, 1.962 Å and 2.000 Å CCSD(T) and SQD respectively. The quantitative differences between SQD and CCSD and CCSD(T) shown in Fig. 1 are a consequence of SQD not being carried out in the full basis set.

Methane Dimer Simulations

Figure 2 shows the binding energy of the methane dimer—with (16e,16o) active space for the dimer and monomer respectively—as a function of the carbon-carbon distance using SQD and HCI. Figure 2 focuses on the attractive region, whereas the full curve is shown in Figure S2 of the Supplementary Note 5. The SQD (16e,16o) and CASCI (16e,16o) data are closely aligned, with deviations below 0.005 kcal/mol. We interpret the excellent agreement between SQD (16e,16o) and CASCI (16e,16o) as an indication that the active-space Schrodinger equation is solved accurately. SQD (16e,16o) predicts the interaction between the monomers to be only marginally attractive, with a binding energy of -0.038 kcal/mol and a lowest-energy distance around 4.500 Å. On the other hand, full-basis CCSD and CCSD(T) calculations predict binding energies of -0.399 kcal/mol and -0.524 kcal/mol, respectively, at distances 3.834 Å and 3.667 Å, respectively. Despite some differences quantifying the importance of perturbative triple corrections, both full-basis calculations predict a substantially more pronounced tendency to binding than SQD (16e,16o) and CASCI (16e,16o). This is because, although SQD (16e,16o) and CASCI (16e,16o) calculations can accurately capture the active-space electronic correlation, they cannot account for the residual dynamical electron correlation, unlike full-basis CCSD and CCSD(T).

Before proceeding with the expansion of the active space we first demonstrate that accurate SQD (16e,16o) calculations can be achieved with a reduced number of samples through the extrapolation of the total energies. The exact SQD (16e,16o) calculations require $|\tilde{\chi}_b| = 20.0 \times 10^3$ while the extrapolation is done based on three points with $|\tilde{\chi}_b|$ of 9.0×10^3, 11.0×10^3, and 14.0×10^3. Hence, the extrapolation allows for the reduction of the maximum required $|\tilde{\chi}_b|$ by 6.0×10^3. We show the SQD (16e,16o) total energy extrapolations for 4.000, 4.250, 4.500, 4.750, 5.000, and 6.000 Å distances in Fig. 3a-f, while the extrapolation for the 48.000 Å distance is shown in Fig. 3g. The resulting binding energies of the methane dimer are shown in 3h and compared against the CASCI (16e,16o) simulations and SQD (16e,16o) simulations with $|\tilde{\chi}_b| = 20.0 \times 10^3$. Figure 3g shows that the extrapolated SQD (16e,16o) energies predict a binding energy in good qualitative agreement with exact SQD (16e,16o) simulations and CASCI (16e,16o). This result is promising for future simulations with large active spaces, where classical post-processing of SQD data becomes computationally expensive.

Methane Dimer Simulations with Extended Active Space

We analyze the effect of extending the active space on the predicted binding energy via the inclusion of virtual orbitals with carbon 3s and 3p character. First, in Fig. 4, we explore the performance of HCI in this extended (16e,24o) active space. Here, HCI is used in place of CASCI due to the fact that the (16e,24o) active space is prohibitively expensive in conventional CASCI simulations. Figure 4a shows active-space calculations with HCI, CCSD, and CCSD(T). The CCSD(T) (16e,24o) curve, in good agreement with CCSD (16e,24o), is substantially more attractive than in the (16e,16o) active space, predicting a binding energy of -0.136 kcal/mol at 4.000 Å. The size of the (16e,24o) active space prevents us

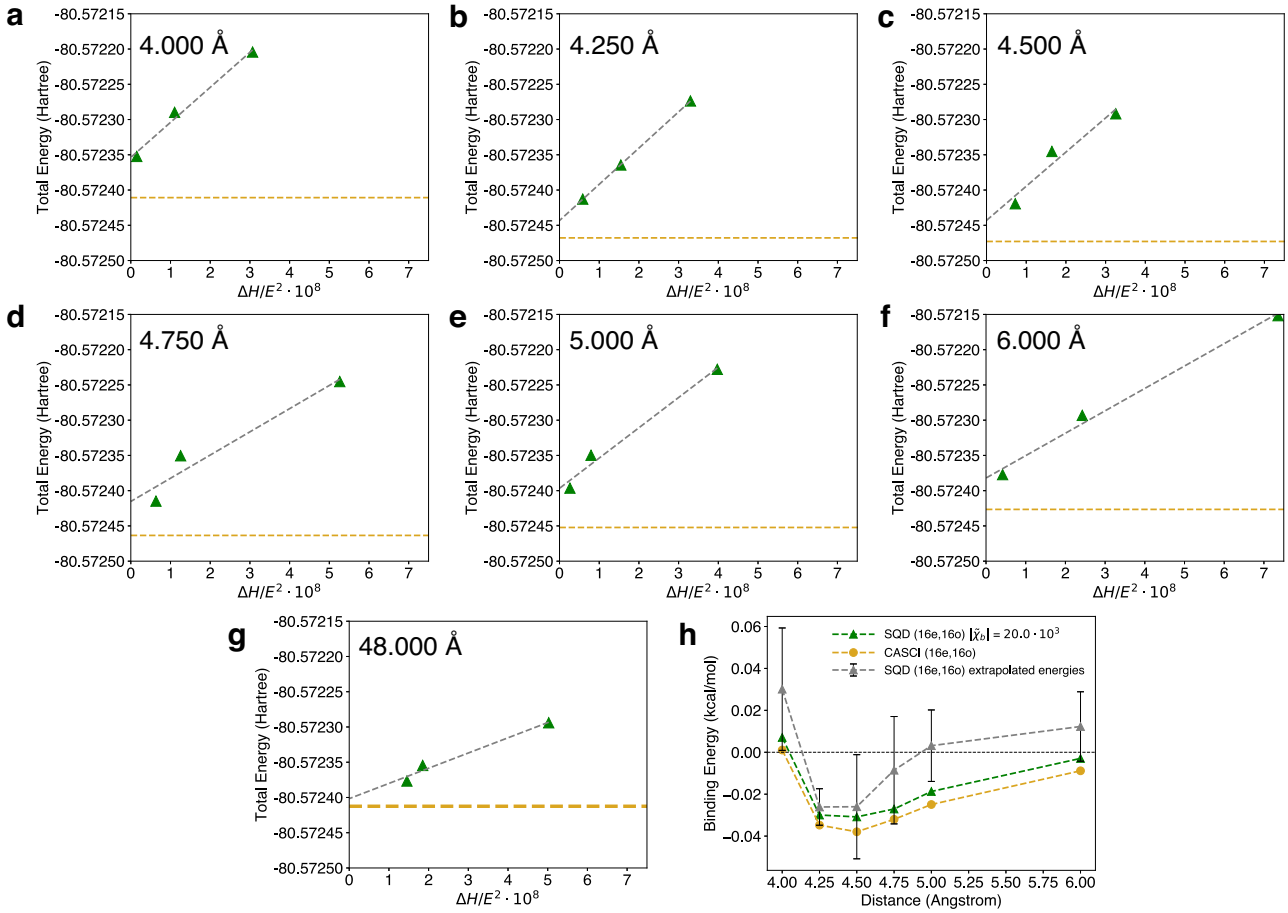

Fig. 3 | Extrapolated SQD (16e,16o) energies of the methane dimer along the PES, for the 4.000, 4.250, 4.500, 4.750, 5.000, and 6.000 Å distances between the carbon atoms. Extrapolations are done using three points with $|\tilde{\chi}_b|$ of 9.0×10^3, 11.0×10^3, and 14.0×10^3. Hamiltonian variance (ΔH) is calculated as $\Delta H = \langle \psi^{(k)} | \hat{H}^2 | \psi^{(k)} \rangle - \langle \psi^{(k)} | \hat{H} | \psi^{(k)} \rangle^2$. **a**–**f** total energy extrapolations for methane dimer 4.000, 4.250, 4.500, 4.750, 5.000, and 6.000 Å distances, **g** total energy extrapolations at 48.000 Å distance, and **h** binding energy in methane dimer calculated with extrapolated SQD (16e,16o) total energies compared against CASCI (16e,16o) simulations and SQD (16e,16o) simulations with $|\tilde{\chi}_b| = 20.0 \times 10^3$. Green triangles and green dashed lines indicate SQD (16e,16o) energies calculated with $|\tilde{\chi}_b| = 20.0 \times 10^3$. Grey triangles and grey dashed lines indicate extrapolated SQD (16e,16o) energies. Light brown circles and dashed lines indicate CASCI (16e,16o) energies. Black horizontal dashed line indicates the zero value of the binding energy. Error bars indicate magnitude of error estimate in extrapolation.

from significantly lowering the parameter ε_1 which results in the underestimation of the total energy in HCI (16e,24o) calculations. In particular, at 3.834 Å distance the HCI (16e,24o) calculations underestimate the binding energy by 0.094 kcal/mol comparing to CCSD(T) (16e,24o), as visible in Fig. 4b. Note that HCI (16e,24o) calculations were carried out over four distances (3.667, 3.750, 3.834, and 3.900 Å). Within the target accuracy of the present work the prediction of HCI (16e,24o) binding energies for other geometries around the CCSD(T) (16e,24o) minimum is dramatically more computationally expensive.

We show the decrease in the SQD (16e,24o) total energy for the methane dimer at 3.638 Å with the increase of $|\tilde{\chi}_b|$ from 5.5×10^3 to $8.5 \cdot 10^3$ in Fig. 5. The differences between the total energies predicted with SQD (16e,24o) and HCI (16e,24o) reduce from 31.8 milliHartree (19.955 kcal/mol) to 25.2 milliHartree (15.813 kcal/mol) when the $|\tilde{\chi}_s|$ is increased from 5.5×10^3 to 8.5×10^3. The extrapolated total energy based on SQD (16e,24o) simulations with $|\tilde{\chi}_b|$ of $5.5 \cdot 10^3$, $6.5 \cdot 10^3$, $7.5 \cdot 10^3$, and 8.5×10^3 agrees with HCI (16e,24o) results within 2.12 milliHartree (1.330 kcal/mol). Magnitude of the error estimate in extrapolation is ± 3.10 milliHartree (1.945 kcal/mol). We believe that a further increase in the number of samples will allow us to advance the accuracy of SQD (16e,24o) calculations. To make SQD (16e,24o) calculations of the methane dimer more computationally feasible we are currently exploring parallelization options for calculations on this system as well as the analysis of the configurations with low contributions to the total energies.

Methodology Limitations and Strategy for Future Improvements

It is generally thought that the first examples of quantum advantage will be achieved for systems with strong electron correlation, for example, in metalloclusters[43], while simulations of non-covalent interactions will be more challenging. The present work showed that the SQD method can treat non-covalent interactions in small molecules with a level of accuracy closely matching the CASCI and HCI methods within the selected active spaces, but did not yet demonstrated quantum advantage in SQD simulations of non-covalent interactions which are dominated by dynamic correlation. In the following we outline exciting future studies that aim to establish the possibility of general quantum advantage in SQD simulations in general and more specifically for non-covalent interactions.

One area of strong interest focuses on the parametrization of the LUCJ circuits. Parametrization of LUCJ circuits based on the t2-amplitudes, obtained from CCSD calculations on classical computers, yields electron configuration distributions corresponding to total energies that differ from the CASCI total energies (see Supplementary Note 6). This necessitates the configuration recovery procedure, which over multiple iterations improves the initial electron configuration distribution[43]. While the configuration recovery procedure yields total energies in close agreement with CASCI total energies, it also increases the cost of the classical computing step due to the need for multiple configuration recovery iterations.

As demonstrated in Supplementary Note 6 showing methane dimer (16e,16o) simulations the optimization of the LUCJ parameters results in

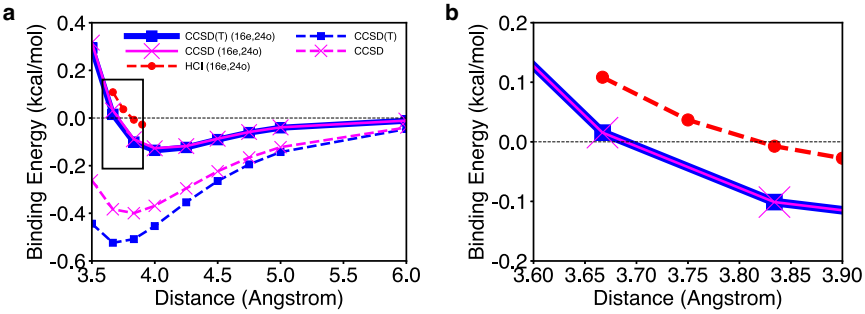

Fig. 4 | Binding energies of the methane dimer along the PES, where the distances between the carbon atoms range between 3.500 and 6.000 Å. Active-space simulations use (16e,24o) and are performed over (**a**) the entire range of bondlengths, and **b** a magnified region near equilibrium, highlighted in (**a**) as a black box. Blue and magenta dashed lines depict PES calculated with CCSD(T) and CCSD methods, respectively. The dashed red line depicts the HCI (16e,24o) results. The solid blue and magenta lines represents CCSD(T) (16e,24o) and CCSD (16e,24o) calculations, respectively. Black horizontal dashed line indicates the zero value of the binding energy.

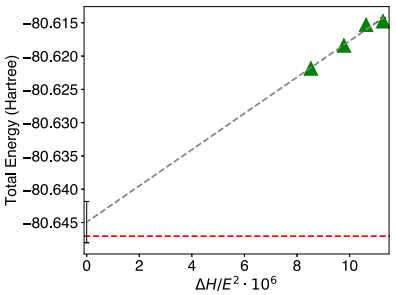

Fig. 5 | Extrapolation of SQD (16e,24o) total energies for the methane dimer at 3.638 Å distance between the carbon atoms. Extrapolations are done using four points with $|\tilde{\chi}_b|$ of $5.5 \cdot 10^3$, $6.5 \cdot 10^3$, $7.5 \cdot 10^3$, and $8.5 \cdot 10^3$. Hamiltonian variance (ΔH) is calculated as $\Delta H = \langle \psi^{(k)}|\hat{H}^2|\psi^{(k)}\rangle - \langle \psi^{(k)}|\hat{H}|\psi^{(k)}\rangle^2$. Green triangles indicate SQD (16e,24o) energies. Grey dashed lines indicate extrapolated SQD (16e,24o) energies. Dashed red line indicates HCI (16e,24o) energies. Error bars indicate magnitude of error estimate in extrapolation.

electron configuration distributions corresponding to total energies that are in much closer agreement with CASCI total energies than the distributions produced with LUCJ circuits parametrized with the t2-amplitudes. Moreover, in the case of noiseless classical simulations of LUCJ circuits with the optimized circuit parameters the resulting distribution yields the total energy of methane dimer (16e,16o) system that is 6.46 milliHartree (4.054 kcal/mol) away from the CASCI energy before use of configuration recovery steps as shown in Figure S3.

Previous studies have shown possible strategies in the optimization of LUCJ circuits[43,73], all of which reduce the resulting LUCJ energy, which shows that optimization procedure for the LUCJ circuit is not unique. Hence, the search for the most optimal strategy is still ongoing and we believe that continuing improvement in LUCJ parameter optimization will reduce the cost of the overall SQD simulations through the reduction of necessary number of configuration recovery steps.

The ability of the LUCJ circuit to generate the electron configurations with high contributions to the lowest energy wavefunction can be further enhanced by increasing the number of density-density interaction terms between the spin-orbitals with opposite spins[46]. In the LUCJ ansatz we only apply density-density interaction terms between qubits that are physically adjacent to each other through a single ancilla qubit. Such an approach allows us to reach the optimal balance between the expressibility of the LUCJ circuits and their depth. The introduction of density-density interaction terms between qubits that are not adjacent to each other would require the introduction of additional SWAP gates[46]. In LUCJ simulations on present day IBM quantum devices utilizing the heavy-hex topology, orbitals for different spins have connections between every 4th orbital (0, 4, 8, etc.), but on quantum computing devices with higher connectivity more density-density interaction terms can be applied in LUCJ circuits without the need to introduce additional SWAP gates. This is another promising direction which can help reach the realization of quantum advantage in simulations with both static and dynamic correlation.

Another focus of ongoing work is on optimizing the post-processing procedure performed on classical computers. As was shown in the SQD paper on the [2Fe-2S] complex[43], noiseless simulations demonstrated that the optimized LUCJ circuits can in principle sample the subspaces more efficiently than the HCI procedure performed classically, where the efficiency of the LUCJ circuits is especially pronounced when compared to the first three iterations of HCI. This advantage of LUCJ is systemically observed across a broad range of HCI ϵ values, where ϵ is the threshold of the Hamiltonian matrix elements used for the enlargement of the subspace in each HCI step. Within the extended SQD framework (ext-SQD)[74] we can extend the sampled subspace also through the use of excitation operators.

Previous ext-SQD studies[74,75] show that extension of the subspaces through single-excitation operators allows for usage of subspaces in the SQD step which are much smaller than the required number of samples in the SQD simulations alone. These results are achieved in runs on quantum hardware even when the underlying LUCJ circuits are parametrized with t2-amplitudes from classical CCSD calculations[74,75]. Moreover, the ext-SQD step is performed after filtering out the SQD electronic configurations that contribute to the overall wavefunction below a given threshold, which means that the expansion of the subspaces in ext-SQD is performed only based on electronic configurations that have the most dominant contributions to lowest energy state. As a result, not only the required SQD subspaces can be significantly reduced, but the final ext-SQD subspace are rather compact[74,75].

Considering that optimized LUCJ circuits[43] are capable of providing an advantage over HCI sampling even when the chosen number of samples is relatively small (which corresponds to the early iterations of HCI) the ext-SQD subspace extension procedure should allow for calculations that overall are less computationally expensive than equivalent HCI calculations. This advantage will be achieved through diagonalization of very compact subspaces (more compact than the ones produced with the HCI procedure) which should substantially reduce the cost of classical compute.

Finally, SQD and ext-SQD calculations can be scaled up for simulations of large systems through embedding schemes. The SQD method allows for simulations of systems with qubit counts essential for the projection-based embedding algorithm proposed by Ralli et al.[76] and can enhance the viability of fragment-based quantum computing simulations. Previously, VQE-based fragment molecular orbital (FMO)[77], divide and conquer (DC)[78], and density matrix embedding theory (DMET)[79,80] simulations were limited to very simple illustrative systems. Shang et al. proposed a DMET-based massively parallel quantum computing approach based on VQE, but execution of their methodology was only possible on a quantum simulator rather than actual hardware[81].

Such limitations in fragment-based VQE simulations are because the number of orbitals that could be described with reasonable accuracy on actual hardware in the VQE formalism within each fragment is very limited. Fragment-based simulations with SQD would allow for substantially higher number of orbitals in each individual fragment, making

Table 1 | Active spaces used in the present work

Species	Active space	AOs	Figure
Water dimer	(16e,12o)	O[2s,2p], H[1s]	6a
Methane dimer	(16e,16o)	C[2s,2p], H[1s]	6b
Methane dimer	(16e,24o)	C[2s,2p,3s,3p], H[1s,2s]	6c

quantum computing simulations of proteins and drug molecules possible. Our preliminary study on DMET simulations with SQD have already showed promise in the prediction of the relative energies of cyclohexane conformers[71]. In the Supplementary Note 7 of the present paper we provide an overview of the DMET SQD methodology and demonstrate that DMET SQD allows for scaling of SQD to (36e,36o) in cyclohexane. We demonstrate in Figure S4 that DMET-SQD can closely reproduce the DMET-FCI total energy. The Hilbert space corresponding to (36e,36o) in cyclohexane is equal to $8.2 \cdot 10^{19}$, which is a few orders of magnitude higher than the $8.9 \cdot 10^{15}$ Hilbert space used for the [4Fe-4S] cluster (54e,36o). It is important to note that even though DMET allows us to reduce the subspaces (via fragmentation) for which we need to perform the diagonalization, the diagonalization itself remains the most expensive step in the SQD procedure.

Conclusions

We have presented quantum-centric simulations of the water and methane dimers using a sample-based quantum diagonalization method on IBM's Eagle quantum processors demonstrating supermolecular treatment of intermolecular interactions on quantum processors. The accuracy of SQD and HCI predictions of non-covalent interactions can be systematically improved by the addition of extended shells of virtual orbitals. We anticipate that further expansion of the active spaces through the inclusion of the virtual orbitals corresponding to the 3d shell of the heavy atoms will allow for an even more accurate description of non-covalent interactions with SQD and HCI, which will be the subject of future studies on quantum processors. Importantly, the present study lays out a framework for SQD electronic structure calculations of non-covalent interactions on quantum hardware.

Our findings demonstrate that SQD is capable of capturing non-covalent interactions between molecules at the chosen level of theory, with potential energy surfaces that closely align with those obtained through classical computational methods. We examine the binding energies of the water and methane dimers by comparing SQD with an analogous classical methods, namely CASCI and HCI. We also compare SQD against the CCSD(T) method, which is considered the gold standard for calculations of binding energies[65]. This comparison aims to evaluate the accuracy of SQD and to understand how the nature of the PES changes with different active-space selections. The ability of HCI and SQD to recover the dispersion interaction is highly dependent on the size and nature of the active space, which is especially critical for predicting the binding energy of the methane dimer. In fact, a previous study by Hapka et al.[82], demonstrated that the ability of intermolecular multi-configurational interaction calculations to recover the dispersion energy depends on the size of the active space and can be improved with systematic expansion of the active space.

The results obtained here demonstrate the improvements both in terms of accuracy and scale of quantum computations on chemical problems, enabling, on current quantum processors, use cases previously thought to belong to the fault-tolerant domain, such as the largest active space considered here for methane, which has 1.3M Pauli operators. Further examples of simulations that could be enabled by our approach include quantum computing simulations of chemical reactivity of CO_2-fixating ruthenium catalyst proposed by Burg et al.[83], Ibrutinib drug simulations proposed by Blunt et al.[84], and the drug-discovery workflows proposed by Pyrkov et al.[85] and Kumar et al.[86], as well as multiple stages of drug optimization as described by Bonde et al.[87].

Finally, we would like to note that while showing promise reducing sampling costs, extrapolation is challenging for complex potential energy surfaces with sharp features or strong non-linearity. We are exploring adaptive sampling strategies, higher-order interpolation methods, and machine-learning techniques to enhance the robustness. Additionally, refining the LUCJ ansatz - such as introducing more connections between alpha and beta spin-orbitals, optimizing ansatz parameters, and integrating hybrid quantum-classical feedback-could enable more efficient configuration selection, further lowering the computational cost of the SQD method. Prior studies suggest these improvements are feasible[43,46]. Moreover, we seek to explore alternative error-mitigation strategies to address noise and decoherence in current and emerging quantum hardware. By leveraging error-mitigated sampling, algorithmic error correction, and more quantum-based solutions, we aim to develop more scalable, hardware-agnostic solutions. As part of this effort we demonstrate in the Supplementary Note 6 how the optimization of the LUCJ parameters allow us to achieve better quantum sampling using the same LUCJ ansatz on quantum hardware.

In conclusion, combining quantum and classical computational resources in workflows like SQD opens the way for the use of current and near-future quantum technology to tackle computational challenges in small-molecule conformational search, drug-protein interactions and drug discovery. The present study lays out the fundamental ground work in simulations of non-covalent interactions with SQD, but to establish the possibility of quantum advantage in simulations of such interactions we anticipate the need for a series of subsequent studies performing extensive screening of various molecular systems to establish specific instances where quantum advantage would be possible. Finally, we would like to note that the demonstration of quantum advantage on real quantum hardware in electronic structure simulations, to our knowledge, has not been achieved by any research group yet, neither with SQD, nor any other quantum-accelerated techniques and, as such, remains one of the most challenging tasks in quantum computing[88]. However, we believe that the SQD-based methodology is one of the promising contenders to reach this goal.

Methods

Classical benchmark

In the supermolecular approach binding energies between two monomers in a dimer is most often expressed as

$$E_{\text{binding}} = E_{\text{AB}} - E_{\text{A}} - E_{\text{B}}. \quad (1)$$

In Eq. (1) E_{AB}, E_{A}, and E_{B} denote the ground-state energies of the dimer *AB*, monomer *A*, and monomer *B*, respectively. For calculations utilizing active spaces the highest accuracy obtainable with the supermolecular approach can be achieved if Eq. (1) is instead expressed in terms of the energy of bound and unbound dimers ($E_{\text{AB-bound}}$ and $E_{\text{AB-unbound}}$). Better accuracy is achieved within this approximation due to the fact that it allows for a consistent active space in all of the calculations. Hence, in all of our calculations we express the binding energy as

$$E_{\text{binding}} = E_{\text{AB}-\text{bound}} - E_{\text{AB}-\text{unbound}}. \quad (2)$$

Here, the $E_{\text{AB-unbound}}$ term of Eq. (2) is approximated as two monomers separated by a 48.000 Å distance, where the chosen distance guarantees the absence of interactions between the monomers.

Metz et al.[89] and Li et al.[90], demonstrated that CCSD(T)/aug-cc-pVQZ calculations closely reproduce the results of the CCSD(T)/complete basis set (CBS) limit for the methane dimer. Metz et al.[89] also demonstrated this for water dimer. All of our simulations are therefore done with the aug-cc-pVQZ basis set. We simulate the water and methane dimers with the active spaces listed in Table 1.

We construct these active spaces using the atomic valence active space (AVAS) method[66] as implemented in the PySCF 2.6.2 software package[68–70], and select active-space orbitals that overlap with the atomic orbitals (AOs) listed in column 3 of the table. The active-space orbitals of the water and

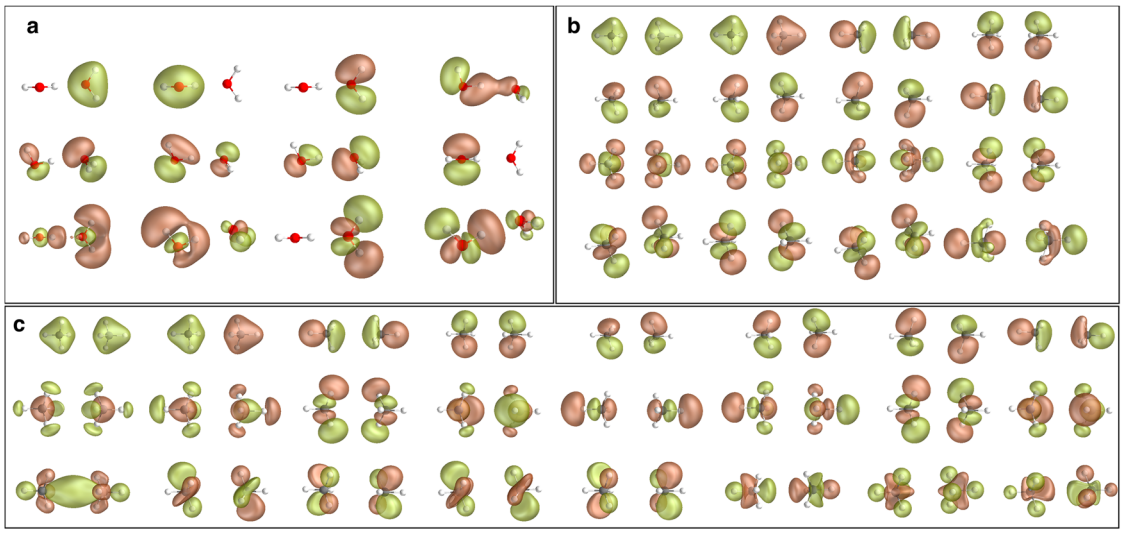

Fig. 6 | Active spaces used in this work. a (16e,12o) of the water dimer, **b** (16e,16o) of the methane dimer, **c** (16e,24o) of the methane dimer.

methane dimers are shown in Fig. 6. Orbital visualization is performed with Pegamoid[91].

In each active space, we perform CCSD and CCSD(T) calculations with PySCF 2.6.2. For water and methane dimers we also perform CAS-CI(16e,12o) and CASCI(16e,16o) simulations, respectively, using PySCF 2.6.2. For the (16e,24o) active space of the methane dimer we perform HCI calculations with the SHCI-SCF 0.1 interface between PySCF 2.6.2 and DICE 1.0[60,62,63]. Further details of HCI calculations can be found in Supplementary Note 2. Along with active-space simulations, we perform complete CCSD and CCSD(T) calculations with ORCA 5.0.4[92]. The geometries of equilibrium structures of the water and methane dimer originate from works by Temelso et al.[93], and by Rezac and Hobza[94], sourced through the BEGDB database[95]. We describe the generation of PES geometries for water and methane dimers in the Supplementary Note 1. The atomic coordinates of the PES structures can be found in Supplementary Data 1.

Quantum Computing

We start from the active-space Hamiltonian, written in second quantization as

$$\hat{H} = E_0 + \sum_{\substack{pr \\ \sigma}} h_{pr} \hat{a}^\dagger_{p\sigma} \hat{a}_{r\sigma} + \sum_{\substack{prqs \\ \sigma\tau}} \frac{(pr|qs)}{2} \hat{a}^\dagger_{p\sigma} \hat{a}^\dagger_{q\tau} \hat{a}_{s\tau} \hat{a}_{r\sigma}, \tag{3}$$

where $\hat{a}^\dagger$ ($\hat{a}$) are creation (annihilation) operators, p,r,s, and $q = 1 \ldots M$ denote basis set element, σ and τ denote spin-z polarizations, h_{pr} and $(pr|qs)$ are the one- and two-body electronic integrals, and E_0 is a constant accounting for the electrostatic interactions between nuclei and electrons in occupied inactive orbitals. We obtain the quantities E_0, h_{pr}, and $(pr|qs)$ for the selected active spaces using PySCF.

We prepare our wavefunction guesses $|\Psi\rangle$, used to approximate the ground state of Eq. (3), from a truncated version of the local unitary cluster Jastrow (LUCJ) ansatz[46]

$$|\Psi\rangle = \prod_{\mu=0}^{L-1} e^{\hat{K}_\mu} e^{i\hat{J}_\mu} e^{-\hat{K}_\mu} |\mathbf{x}_{\mathrm{RHF}}\rangle, \tag{4}$$

where $\hat{K}_\mu = \Sigma_{pr,\sigma} K^\mu_{pr} \hat{a}^\dagger_{p\sigma} \hat{a}_{r\sigma}$ are one-body operators, $\hat{J}_\mu = \Sigma_{pr,\sigma\tau} J^\mu_{p\sigma,r\tau} \hat{n}_{p\sigma} \hat{n}_{r\tau}$ are suitable (vide infra) density-density operators, and $|\mathbf{x}_{\mathrm{RHF}}\rangle$ is the restricted closed-shell Hartree-Fock (RHF) state. We use the Jordan-Wigner (JW) transformation[96] to map the fermionic wavefunction Eq. (4) onto a qubit wavefunction that can be prepared executing a quantum circuit. The JW transformation maps the Fock space of fermions in M spatial orbitals onto the Hilbert space of $2M$ qubits, where the basis state $|\mathbf{x}\rangle$ is parametrized by a bitstring $\mathbf{x} \in \{0, 1\}^{2M}$ and represents an electronic configuration where the spin-orbital $p\sigma$ is occupied (empty) if $x_{p\sigma} = 1$ ($x_{p\sigma} = 0$).The details of the JW transformation can be found in the Supplementary Note 3.

We prepare the wavefunction Eq. (4) by executing the following quantum circuit: a single layer of Pauli-X gates prepares the basis state $|\mathbf{x}_{\mathrm{RHF}}\rangle$, a Bogoliubov circuit[97] (with linear depth, quadratic number of gates, and a 1D qubit connectivity) encodes each orbital rotation $e^{\pm \hat{K}_\mu}$, and a circuit of Pauli-ZZ rotations encodes each density-density interaction $e^{i\hat{J}_\mu}$. When J^μ is a dense matrix, Pauli-ZZ rotations are applied across all pair of qubits, requiring all-to-all qubit connectivity or a substantial overhead of swap gates. To mitigate these quantum hardware requirements LUCJ imposes a "locality" approximation, i.e., it assumes $J^\mu_{p\sigma,r\tau} = 0$ for all pairs of spin-orbitals that are not mapped onto adjacent qubits under JW[46] (as a consequence, a circuit with constant depth and linear number of gates encodes each $e^{i\hat{J}_\mu}$ operator). Hence, the number of layers $(L-1)$ in Eq. (4) is formally equal to 1.5. As the result the specific form of $|\Psi\rangle$ used in this work is expressed as $|\Psi\rangle = e^{-\hat{K}_2} e^{\hat{K}_1} e^{i\hat{J}_1} e^{-\hat{K}_1} |\mathbf{x}_{\mathrm{RHF}}\rangle$. We parametrize the LUCJ circuit based on amplitudes computed from classical restricted closed-shell CCSD within the given active space[43], yet a further quantum-classical parameter optimization could further improve the quality of the ground-state approximation. We produce the LUCJ circuits using the ffsim library[98] interfaced with Qiskit 1.1.1[97,99].

The qubit layouts of the LUCJ circuits used for (16e,12o) water dimer, (16e,16o) methane dimer, and (16e,24o) methane dimer simulations are shown in Figs. 7a, b, and c, respectively. We execute these circuits on IBM's 127-qubit Eagle devices ibm_cleveland and ibm_kyiv. In all our quantum computing experiments, we used gate (not measurement) twirling over random 2-qubit Clifford gates[100] and dynamical decoupling[101–104] – available through the SamplerV2 primitive of Qiskit's runtime library – to mitigate quantum errors. The number of qubits, 2-qubit gate depth, and number of CNOT gates in the LUCJ circuits are shown in Fig. 7d.

Upon executing the LUCJ circuits, we measure $|\Psi\rangle$ in the computational basis. Repeating this produces a set of measurement outcomes (or "shots")

$$\tilde{\chi} = \{\mathbf{x} | \mathbf{x} \sim \tilde{p}(\mathbf{x})\} \tag{5}$$

in the form of bitstrings $\mathbf{x} \in \{0, 1\}^{2M}$, each representing an electronic configuration (Slater determinants) distributed according to $\tilde{p}(\mathbf{x})$. While on a noiseless device configurations are distributed according to $|\langle \mathbf{x}|\Psi\rangle|^2$, on a

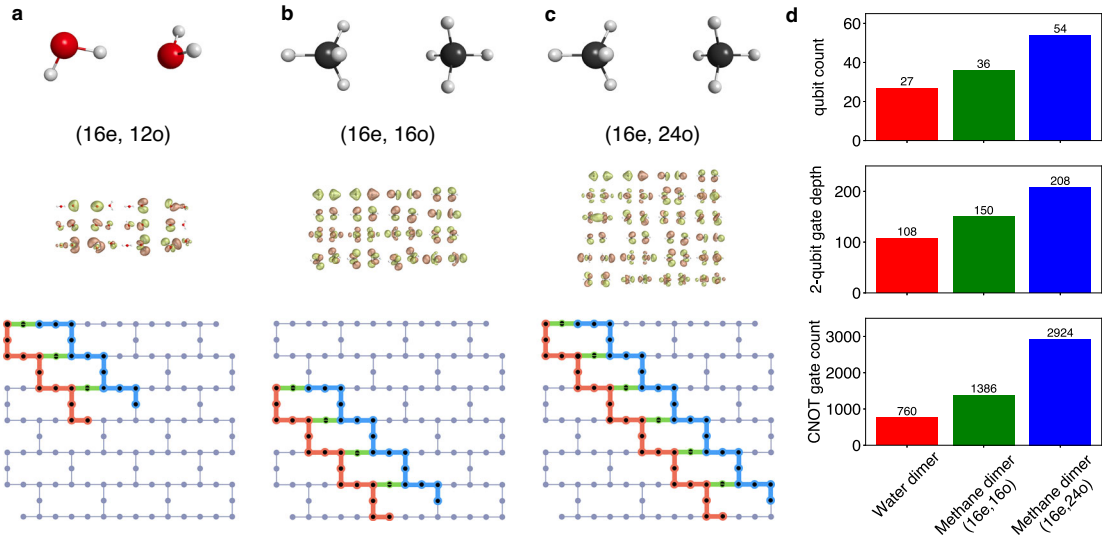

Fig. 7 | Overview of quantum circuits. a–c Qubit layouts of LUCJ circuits executed in this work: **a** (16e,12o) water dimer simulations using 27 qubits on ibm_cleveland, **b** (16e,16o) methane dimer simulations using 36 qubits on ibm_cleveland, and **c** (16e,24o) methane dimer simulations using 54 qubits on ibm_kyiv. **d** Number of qubits, 2-qubit gate depth, and CNOT gate count of the LUCJ circuits.Qubits used to encode occupation numbers of α (β) spin-orbitals are shown in red (blue). Auxiliary qubits used to execute density-density interactions between α and β spin-orbitals are marked in green. Structures and molecular orbitals of corresponding systems are represented above the qubit layouts.

Table 2 | Details of SQD calculations

active space	$\lvert\tilde{\chi}\rvert$ [10^3]	K	$\lvert\tilde{\chi}_b\rvert$ [10^3]	d	Hilbert space	CPUs, code	steps
(16e,12o)	200	10	0.8	$17.2 \cdot 10^4$	$24.5 \cdot 10^4$	10, PySCF	10
(16e,16o)	200	10	20.0	$12.6 \cdot 10^7$	$16.6 \cdot 10^7$	10, PySCF	10
(16e,24o)	300	4	8.5	$24.9 \cdot 10^7$	$54.1 \cdot 10^{10}$	16, DICE	5

noisy device they follow a distribution $\tilde{p}(\mathbf{x}) \neq |\langle \mathbf{x}|\Psi\rangle|^2$. In particular, $\tilde{p}(\mathbf{x})$ breaks particle-number conservation and returns configurations with incorrect particle number. We use a technique called self-consistent configuration recovery[43], executed on a classical computer, to restore particle-number conservation. The associated code is publicly available in the GitHub repository[67]. Within each step of self-consistent recovery, we sample K subsets (or batches) of $\tilde{\chi}$ labeled $\tilde{\chi}_b$ with $b = 1 \ldots K$. Each batch defines – through a transformation[43] informed by an approximation to the ground-state occupation numbers $n_{p\sigma}$ – a subspace $S^{(b)}$ of dimension d, in which we project the many-electron Hamiltonian, using classical computational resources, as[43,44,105]

$$\hat{H}_{S^{(b)}} = \hat{P}_{S^{(b)}} \hat{H} \hat{P}_{S^{(b)}}, \tag{6}$$

where the projector $\hat{P}_{S^{(b)}}$ is

$$\hat{P}_{S^{(b)}} = \sum_{\mathbf{x} \in S^{(b)}} |\mathbf{x}\rangle\langle\mathbf{x}|. \tag{7}$$

The subspaces $S^{(b)}$ constructed with extension of the set of configurations $\tilde{\chi}_b$ to ensure the closure under spin inversion symmetry[43], which produces the larger values of d than $\tilde{\chi}_b$. We compute the ground states and energies of the Hamiltonians in Eq. (6), $|\psi^{(b)}\rangle$ and $E^{(b)}$ respectively, and use the lowest energy across the batches, $\min_b E^{(b)}$, as the best approximation to the ground-state energy at the current iteration of the configuration recovery. We use the ground states $|\psi^{(b)}\rangle$ to obtain an updated set of occupation numbers,

$$n_{p\sigma} = \frac{1}{K} \sum_{1 \le b \le K} \langle \psi^{(b)} | \hat{n}_{p\sigma} | \psi^{(b)} \rangle, \tag{8}$$

that we use in the next iteration of configuration recovery to produce the subspaces $S^{(b)}$. We repeat the iterations of self-consistent configuration recovery until convergence of the energy $\min_b E^{(b)}$. At the first iteration of self-consistent configuration recovery, we initialize $n_{p\sigma}$ from the measurement outcomes in $\tilde{\chi}$ with the correct particle number. We summarize the details of our SQD calculations in Table 2.

We demonstrate that for SQD (16e,16o) simulations of the methane dimer at 3.638 Å a $|\tilde{\chi}_b| = 20.0 \cdot 10^3$ is necessary to reach agreement within 0.010 kcal/mol when compared against CASCI (16e,16o). We show how the predicted total energies in these simulations improve with an increase of $|\tilde{\chi}_b|$ from $5.0 \cdot 10^3$ to $20.0 \cdot 10^3$ in Figure S1 of Supplementary Note 4. We also demonstrate that in SQD(16e,16o) simulations the linear energy-variance relation allows for utilization of energy extrapolation which reproduces similar binding energies as simulation with $|\tilde{\chi}_b| = 20.0 \cdot 10^3$ while using substantially lower values of $|\tilde{\chi}_b|$. The extrapolation is done for the total energy of the dimer as the function of the Hamiltonian variance divided by the square of the variational energy, where Hamiltonian variance (ΔH) is calculated as $\Delta H = \langle \psi^{(k)} | \hat{H}^2 | \psi^{(k)} \rangle - \langle \psi^{(k)} | \hat{H} | \psi^{(k)} \rangle^2$. The extrapolation is done based on three points with $|\tilde{\chi}_b|$ of $9.0 \cdot 10^3$, $11.0 \cdot 10^3$, and $14.0 \cdot 10^3$, which allows for the reduction of the maximum $|\tilde{\chi}_b|$ by $6.0 \cdot 10^3$. This choice of values for $|\tilde{\chi}_b|$ allows for an even distribution of ΔH values used in extrapolation. The extrapolated energies are compared against CASCI(16e,16o) simulations and SQD(16e,16o) simulations with $|\tilde{\chi}_b| = 20.0 \cdot 10^3$.

We compute the ground-state eigenpairs of the Hamiltonians Eq. (6) using the iterative Davidson method on 10 CPUs with PySCF's selected configuration interaction (SCI) solver for SQD (16e,12o) simulations of the water dimer and SQD (16e,16o) simulations of the methane dimer. We achieve parallelization across 10 CPUs with Ray 2.33.0[106] where the eigenstate solver within each of the 10 batches is using 1 CPU. For SQD (16e,24o) simulations of the methane dimer, we utilize the SCI solver of DICE and 16

CPUs, where the eigenstate solver within each of the 4 batches is using 4 CPUs. Further parallelization is possible with the SCI solver of DICE, as was demonstrated previously[43]. The SQD (16e,12o) simulations of the water dimer and SQD (16e,16o) simulations of the methane dimer are done for the distances that are described in the Supplemental Note 1 while SQD (16e,24o) simulations of the methane dimer are only done for 3.638 Å.

Data availability

The equilibrium structures of water and methane dimers are available through BEGDB entrances below: http://www.begdb.org/index.php?action=oneMolecule&state=show&id=4179 http://www.begdb.org/index.php?action=oneMolecule&state=show&id=3975. All other relevant data are available from the authors upon the request.

Code availability

Qiskit, ffsim, and Qiskit IBM Runtime utilized for LUCJ simulations can be obtained from corresponding GitHub repositories: https://github.com/qiskit-community/ffsim https://github.com/Qiskit/qiskit https://github.com/Qiskit/qiskit-ibm-runtime. Configuration recovery code is distributed as Qiskit SQD addon and includes the option for parallel DICE-based solver: https://github.com/Qiskit/qiskit-addon-sqd https://github.com/Qiskit/qiskit-addon-dice-solver. PySCF and its interface with DICE are available through corresponding GitHub repositories: https://github.com/pyscf/pyscf https://github.com/pyscf/shciscf. The tutorial demonstrating full SQD workflow is available below: https://qiskit.github.io/qiskit-addon-sqd/tutorials/01_chemistry_hamiltonian.html.

Received: 13 December 2024; Accepted: 2 September 2025;
Published online: 08 October 2025

References

1. Müller-Dethlefs, K. & Hobza, P. Noncovalent interactions: a challenge for experiment and theory. *Chemical Reviews* **100**, 143–168 (2000).
2. Puzzarini, C., Spada, L., Alessandrini, S. & Barone, V. The challenge of non-covalent interactions: Theory meets experiment for reconciling accuracy and interpretation. *Journal of Physics: Condensed Matter* **32**, 343002 (2020).
3. Pichler, W. J. The important role of non-covalent drug-protein interactions in drug hypersensitivity reactions. *Allergy* **77**, 404–415 (2022).
4. Aljoundi, A., Bjij, I., El Rashedy, A. & Soliman, M. E. Covalent versus non-covalent enzyme inhibition: which route should we take? a justification of the good and bad from molecular modelling perspective. *The Protein Journal* **39**, 97–105 (2020).
5. Dyson, H. J., Wright, P. E. & Scheraga, H. A. The role of hydrophobic interactions in initiation and propagation of protein folding. *Proceedings of the National Academy of Sciences* **103**, 13057–13061 (2006).
6. Zondlo, N. J. Fold globally, bond locally. *Nature Chemical Biology* **6**, 567–568 (2010).
7. Pace, C. N. et al. Contribution of hydrophobic interactions to protein stability. *J Mol Biol* **408**, 514–28 (2011).
8. Camilloni, C. et al. Towards a structural biology of the hydrophobic effect in protein folding. *Scientific Reports* **6**, 28285 (2016).
9. Durell, S. R. & Ben-Naim, A. Hydrophobic-hydrophilic forces in protein folding. *Biopolymers* **107** (2017).
10. Hanshaw, R. G., Stahelin, R. V. & Smith, B. D. Noncovalent keystone interactions controlling biomembrane structure. *Chemistry* **14**, 1690–7 (2008).
11. Adhav, V. A. & Saikrishnan, K. The realm of unconventional noncovalent interactions in proteins: Their significance in structure and function. *ACS Omega* **8**, 22268–22284 (2023).
12. Jhoti, H. & Leach, A. R.*Structure-Based Drug Discovery* (Springer Netherlands, 2007).
13. Gros, E. et al. A non-covalent peptide-based strategy for protein and peptide nucleic acid transduction. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1758**, 384–393 (2006).
14. Wei, Y., Ma, L., Zhang, L. & Xu, X. Noncovalent interaction-assisted drug delivery system with highly efficient uptake and release of paclitaxel for anticancer therapy. *Int J Nanomedicine* **12**, 7039–7051 (2017).
15. Yang, Y. et al. Effect of hydrophobicity on the anticancer activity of fatty-acyl-conjugated cm4 in breast cancer cells. *ACS Omega* **5**, 21513–21523 (2020).
16. Wright, K. M. et al. Hydrophobic interactions dominate the recognition of a kras g12v neoantigen. *Nat Commun* **14**, 5063 (2023).
17. Riley, K. E. & Hobza, P. Noncovalent interactions in biochemistry. *WIREs Computational Molecular Science* **1**, 3–17 (2011).
18. DiLabio, G. A. & Otero-de-la Roza, A. Noncovalent interactions in density functional theory. *Reviews in computational chemistry* **29**, 1–97 (2016).
19. Burns, L. A., Mayagoitia, A. V., Sumpter, B. G. & Sherrill, C. D. Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals. *The Journal of Chemical Physics* **134**, 084107 (2011).
20. Gao, H., Imamura, S., Kasagi, A. & Yoshida, E. Distributed implementation of full configuration interaction for one trillion determinants. *Journal of Chemical Theory and Computation* **20**, 1185–1192 (2024).
21. Vogiatzis, K. D., Ma, D., Olsen, J., Gagliardi, L. & de Jong, W. A. Pushing configuration-interaction to the limit: Towards massively parallel MCSCF calculations. *The Journal of Chemical Physics* **147**, 184111 (2017).
22. Menczer, A. et al. Parallel implementation of the density matrix renormalization group method achieving a quarter petaflops performance on a single dgx-h100 gpu node. *Journal of Chemical Theory and Computation* **0**, null (2024).
23. Ren, X., Zou, J., Zhang, H., Li, W. & Li, S. Block-correlated coupled cluster theory with up to four-pair correlation for accurate static correlation of strongly correlated systems. *The Journal of Physical Chemistry Letters* **15**, 693–700 (2024).
24. Jin, R. et al. PASCI: A scalable framework for heterogeneous parallel calculation of dynamical electron correlation (2024).
25. Schwenke, D. W. Introducing MPEC: Massively parallel electron correlation. *The Journal of Chemical Physics* **158**, 084801 (2023).
26. Delcey, M. G. Multipsi: A python-driven mcscf program for photochemistry and spectroscopy simulations on modern hpc environments. *WIREs Computational Molecular Science* **13**, e1675 (2023).
27. Corzo, H. H. et al. Coupled cluster theory on modern heterogeneous supercomputers. *Frontiers in Chemistry***11** (2023).
28. Datta, D. & Gordon, M. S. Accelerating coupled-cluster calculations with gpus: An implementation of the density-fitted ccsd(t) approach for heterogeneous computing architectures using openmp directives. *Journal of Chemical Theory and Computation* **19**, 7640–7657 (2023).
29. Zhai, H. & Chan, G. K.-L. Low communication high performance ab initio density matrix renormalization group algorithms. *The Journal of Chemical Physics* **154**, 224116 (2021).
30. He, X., Walker, B., Man, V. H., Ren, P. & Wang, J. Recent progress in general force fields of small molecules. *Curr Opin Struct Biol* **72**, 187–193 (2022).
31. Nerenberg, P. S. & Head-Gordon, T. New developments in force fields for biomolecular simulations. *Current Opinion in Structural Biology* **49**, 129–138 (2018).
32. Ponder, J. W. & Case, D. A. Force fields for protein simulations. *Adv. Protein Chem.* **66**, 27–85 (2003).

33. Mackerell, J. A. D. Empirical force fields for biological macromolecules: overview and issues. *J Comput Chem* **25**, 1584–604 (2004).
34. Jorgensen, W. L. Monte carlo simulations for free energies of hydration: Past to present. *J Chem Phys* **161** (2024).
35. Unke, O. T. et al. Machine learning force fields. *Chemical Reviews* **121**, 10142–10186 (2021).
36. Smith, J. S., Isayev, O. & Roitberg, A. E. Ani-1: an extensible neural network potential with dft accuracy at force field computational cost. *Chem. Sci.* **8**, 3192–3203 (2017).
37. Rezaee, M., Ekrami, S. & Hashemianzadeh, S. M. Comparing ani-2x, ani-1ccx neural networks, force field, and dft methods for predicting conformational potential energy of organic molecules. *Scientific Reports* **14**, 11791 (2024).
38. Temme, K., Osborne, T. J., Vollbrecht, K. G., Poulin, D. & Verstraete, F. Quantum metropolis sampling. *Nature* **471**, 87–90 (2011).
39. Yung, M.-H. & Aspuru-Guzik, A. A quantum–quantum metropolis algorithm. *Proceedings of the National Academy of Sciences* **109**, 754–759 (2012).
40. Davoudi, Z., Mueller, N. & Powers, C. Towards quantum computing phase diagrams of gauge theories with thermal pure quantum states. *Physical Review Letters* **131**, 081901 (2023).
41. Watts, O., Kikuchi, Y. & Coopmans, L. Quantum semidefinite programming with thermal pure quantum states. *arXiv preprint arXiv:2310.07774* (2023).
42. Alexeev, Y. et al. Quantum-centric supercomputing for materials science: A perspective on challenges and future directions. *Future Generation Computer Systems* **160**, 666–710 (2024).
43. Robledo-Moreno, J. et al. Chemistry beyond the scale of exact diagonalization on a quantum-centric supercomputer. *Science Advances* **11**, eadu9991 (2025).
44. Kanno, K. et al. Quantum-selected configuration interaction: Classical diagonalization of hamiltonians in subspaces selected by quantum computers. *arXiv preprint arXiv:2302.11320* (2023).
45. Anand, A. et al. A quantum computing view on unitary coupled cluster theory. *Chemical Society Reviews* **51**, 1659–1684 (2022).
46. Motta, M., Sung, K. J., Whaley, K. B., Head-Gordon, M. & Shee, J. Bridging physical intuition and hardware efficiency for correlated electronic states: the local unitary cluster jastrow ansatz for electronic structure. *Chemical Science* **14**, 11213–11227 (2023).
47. Nguyen, B. D. et al. Divergence of many-body perturbation theory for noncovalent interactions of large molecules. *Journal of Chemical Theory and Computation* **16**, 2258–2273 (2020).
48. Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491–495 (2019).
49. Google AI Quantum & Collaborators Hartree-fock on a superconducting qubit quantum computer. *Science* **369**, 1084–1089 (2020).
50. Grimsley, H. R., Economou, S. E., Barnes, E. & Mayhall, N. J. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature Communications* **10**, 3007 (2019).
51. Eddins, A. et al. Doubling the size of quantum simulators by entanglement forging. *PRX Quantum* **3**, 010309 (2022).
52. Liepuoniute, I. et al. Simulation of a diels–alder reaction on a quantum computer. *Physical Chemistry Chemical Physics* **26**, 25181–25191 (2024).
53. Weaving, T., Ralli, A., Love, P. J., Succi, S. & Coveney, P. V. Contextual subspace variational quantum eigensolver calculation of the dissociation curve of molecular nitrogen on a superconducting quantum computer. *npj Quantum Information* **11**, 25 (2025).
54. Loipersberger, M. et al. Accurate non-covalent interaction energies on noisy intermediate-scale quantum computers via second-order symmetry-adapted perturbation theory. *Chemical Science* **14**, 3587–3599 (2023).
55. Ollitrault, P. J. et al. Estimation of electrostatic interaction energies on a trapped-ion quantum computer. *ACS Central Science* **10**, 882–889 (2024).
56. Jeziorski, B., Moszynski, R. & Szalewicz, K. Perturbation theory approach to intermolecular potential energy surfaces of van der waals complexes. *Chemical Reviews* **94**, 1887–1930 (1994).
57. Jeziorski, B., Bulski, M. & Piela, L. First-order perturbation treatment of the short-range repulsion in a system of many closed-shell atoms or molecules. *International Journal of Quantum Chemistry* **10**, 281–297 (1976).
58. Patkowski, K. Recent developments in symmetry-adapted perturbation theory. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **10**, e1452 (2020).
59. Anderson, L. W. et al. Coarse-grained intermolecular interactions on quantum processors. *Physical Review A* **105**, 062409 (2022).
60. Holmes, A. A., Tubman, N. M. & Umrigar, C. Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling. *Journal of Chemical Theory and Computation* **12**, 3674–3680 (2016).
61. Holmes, A. A., Changlani, H. J. & Umrigar, C. Efficient heat-bath sampling in fock space. *Journal of Chemical Theory and Computation* **12**, 1561–1571 (2016).
62. Smith, J. E., Mussard, B., Holmes, A. A. & Sharma, S. Cheap and near exact casscf with large active spaces. *Journal of Chemical Theory and Computation* **13**, 5468–5478 (2017).
63. Sharma, S., Holmes, A. A., Jeanmairet, G., Alavi, A. & Umrigar, C. J. Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory. *Journal of Chemical Theory and Computation* **13**, 1595–1604 (2017).
64. Bartlett, R. J. Perspective on coupled-cluster theory. the evolution toward simplicity in quantum chemistry. *Physical Chemistry Chemical Physics* **26**, 8013–8037 (2024).
65. Cársky, P., Paldus, J. & Pittner, J.*Recent progress in coupled cluster methods: Theory and applications*, vol. 10 of *Challenges and Advances in Computational Chemistry and Physics* (Springer Science & Business Media, Dordrecht, 2010).
66. Sayfutyarova, E. R., Sun, Q., Chan, G. K.-L. & Knizia, G. Automated construction of molecular active spaces from atomic valence orbitals. *Journal of Chemical Theory and Computation* **13**, 4063–4078 (2017).
67. Saki, A. A. et al. Qiskit addon: sample-based quantum diagonalization. https://github.com/Qiskit/qiskit-addon-sqd (2024).
68. Sun, Q. et al. Recent developments in the pyscf program package. *The Journal of Chemical Physics* **153**, 024109 (2020).
69. Sun, Q. et al. Pyscf: the python-based simulations of chemistry framework. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **8**, e1340 (2018).
70. Sun, Q. Libcint: An efficient general integral library for gaussian basis functions. *Journal of Computational Chemistry* **36**, 1664–1671 (2015).
71. Shajan, A. et al. Toward quantum-centric simulations of extended molecules: Sample-based quantum diagonalization enhanced with density matrix embedding theory. *Journal of Chemical Theory and Computation* **21**, 6801–6810 (2025).
72. Kaliakin, D., Shajan, A., Liang, F. & Merz, K. M. J. Implicit solvent sample-based quantum diagonalization. *The Journal of Physical Chemistry B* **129**, 5788–5796 (2025).
73. Motta, M., Sung, K. J. & Shee, J. Quantum algorithms for the variational optimization of correlated electronic states with stochastic reconfiguration and the linear method. *The Journal of Physical Chemistry A* **128**, 8762–8776 (2024).
74. Barison, S., Moreno, J. R. & Motta, M. Quantum-centric computation of molecular excited states with extended sample-based quantum

diagonalization. *Quantum Science and Technology* **10**, 025034 (2025).
75. Barroca, M. A. et al. Surface reaction simulations for battery materials through sample-based quantum diagonalization and local embedding. *arXiv preprint arXiv:2503.10923* (2025).
76. Ralli, A., Williams de la Bastida, M. & Coveney, P. V. Scalable approach to quantum simulation via projection-based embedding. *Phys. Rev. A* **109**, 022418 (2024).
77. Lim, H. et al. Fragment molecular orbital-based variational quantum eigensolver for quantum chemistry in the age of quantum computing. *Scientific Reports* **14**, 2422 (2024).
78. Yoshikawa, T., Takanashi, T. & Nakai, H. Quantum algorithm of the divide-and-conquer unitary coupled cluster method with a variational quantum eigensolver. *Journal of Chemical Theory and Computation* **18**, 5360–5373 (2022).
79. Kawashima, Y. et al. Optimizing electronic structure simulations on a trapped-ion quantum computer using problem decomposition. *Communications Physics* **4**, 245 (2021).
80. Iijima, N. et al. Towards accurate quantum chemical calculations on noisy quantum computers. *arXiv preprint arXiv:2311.09634* (2023).
81. Shang, H. et al. Towards practical and massively parallel quantum computing emulation for quantum chemistry. *npj Quantum Information* **9**, 33 (2023).
82. Hapka, M., Krzemińska, A. & Pernal, K. How much dispersion energy is included in the multiconfigurational interaction energy? *Journal of Chemical Theory and Computation* **16**, 6280–6293 (2020).
83. von Burg, V. et al. Quantum computing enhanced computational catalysis. *Physical Review Research* **3**, 033055 (2021).
84. Blunt, N. S. et al. Perspective on the current state-of-the-art of quantum computing for drug discovery applications. *Journal of Chemical Theory and Computation* **18**, 7001–7023 (2022).
85. Pyrkov, A. et al. Quantum computing for near-term applications in generative chemistry and drug discovery. *Drug Discovery Today* **28**, 103675 (2023).
86. Kumar, G., Yadav, S., Mukherjee, A., Hassija, V. & Guizani, M. Recent advances in quantum computing for drug discovery and development. *IEEE Access* **12**, 64491–64509 (2024).
87. Bhushan, B., Pratik, P. & Bhaskar, C. Chapter 7: The future of drug development with quantum computing. In Heifetz, A. (ed.) *High Performance Computing for Drug Discovery and Biomedicine*, 153–179 (Springer, 2023).
88. Lee, S. et al. Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry. *Nature communications* **14**, 1952 (2023).
89. Metz, M. P. et al. Molecular dimers of methane clathrates: ab initio potential energy surfaces and variational vibrational states. *Physical Chemistry Chemical Physics* **21**, 13504–13525 (2019).
90. Li, A. H.-T. & Chao, S. D. Interaction energies of dispersion-bound methane dimer from coupled cluster method at complete basis set limit. *Journal of Molecular Structure: THEOCHEM* **897**, 90–94 (2009).
91. Galván, I. F. Pegamoid https://gitlab.com/Jellby/Pegamoid (2024).
92. Neese, F. Software update: The orca program system - version 5.0. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **12**, e1606 (2022).
93. Temelso, B., Archer, K. A. & Shields, G. C. Benchmark structures and binding energies of small water clusters with anharmonicity corrections. *The Journal of Physical Chemistry A* **115**, 12034–12046 (2011).
94. Rezac, J. & Hobza, P. Describing noncovalent interactions beyond the common approximations: how accurate is the “gold standard,” ccsd(t) at the complete basis set limit? *Journal of Chemical Theory and Computation* **9**, 2151–2155 (2013).
95. R^ezáč, J. et al. Quantum chemical benchmark energy and geometry database for molecular clusters and complex molecular systems (https://www.begdb.com): a user’s manual and examples. *Collection of Czechoslovak Chemical Communications* **73**, 1261–1270 (2008).
96. Jordan, P. & Wigner, E. P.*Über das paulische Äquivalenzverbot* (Springer, 1993).
97. Aleksandrowicz, G. et al. Qiskit: An open-source framework for quantum computing (2019).
98. The ffsim developers. ffsim: Faster simulations of fermionic quantum circuits. https://github.com/qiskit-community/ffsim (2024).
99. Javadi-Abhari, A. et al. Quantum computing with qiskit. *arXiv preprint arXiv:2405.08810* (2024).
100. Wallman, J. J. & Emerson, J. Noise tailoring for scalable quantum computation via randomized compiling. *Physical Review A* **94**, 052325 (2016).
101. Viola, L. & Lloyd, S. Dynamical suppression of decoherence in two-state quantum systems. *Physical Review A* **58**, 2733 (1998).
102. Kofman, A. & Kurizki, G. Universal dynamical control of quantum mechanical decay: modulation of the coupling to the continuum. *Physical Review Letters* **87**, 270405 (2001).
103. Biercuk, M. J. et al. Optimized dynamical decoupling in a model quantum memory. *Nature* **458**, 996–1000 (2009).
104. Niu, S. & Todri-Sanial, A. Effects of dynamical decoupling and pulse-level optimizations on ibm quantum computers. *IEEE Transactions on Quantum Engineering* **3**, 1–10 (2022).
105. Nakagawa, Y. O., Kamoshita, M., Mizukami, W., Sudo, S. & Ohnishi, Y.-y Adapt-qsci: Adaptive construction of an input state for quantum-selected configuration interaction. *Journal of Chemical Theory and Computation* **20**, 10817–10825 (2024).
106. Moritz, P. et al. *Ray: A Distributed Framework for Emerging AI Applications* (USENIX Association, Carlsbad, CA, 2018). https://www.usenix.org/conference/osdi18/presentation/moritz.

Acknowledgements

We gratefully acknowledge financial support from the National Science Foundation (NSF) through CSSI Frameworks Grant OAC-2209717 and from the National Institutes of Health (Grant Numbers GM130641).

Author contributions

D.K., A.S., and F.L. planned the simulations; performed the simulations; analyzed the results; have written the manuscript. S.D. performed the simulations; analyzed the results; have written the manuscript. A.M. assisted with the simulations planning; have written the manuscript. J.R.M. and M.M. developed the methodology; provided the guidance on methodology to the entire team; performed the software development; planned the simulations; performed the simulations; have written the manuscript. Z.L. addressed the parallelization and debugging of the code; performed the simulations; analyzed the results; have written the manuscript. C.J. performed code development; addressed the parallelization and debugging; provided the guidance on the code development and optimization of the code to the entire team. I.S. performed the code development; provided guidance on software stack and optimal usage of classical and quantum hardware. A.A.S. helped to optimize the efficiency of classical computing in configuration recovery procedure; provided the guidance on efficient quantum computing simulations in Qiskit and optimal ways to mitigate quantum error. A.M. and K.M.M. supervised the entire project, planned the experiments, have written the manuscript.

Competing interests

The authors declare no competing interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42005-025-02305-9.

Correspondence and requests for materials should be addressed to Kenneth M. Merz Jr.

Peer review information *Communications Physics* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.