Quantum-centric computation of molecular excited states with extended sample-based quantum diagonalization

Stefano Barison, 1, 2, 3, * Javier Robledo Moreno, 4, † and Mario Motta^{4, ‡}

¹Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
²Center for Quantum Science and Engineering, EPFL, Lausanne, Switzerland
³National Centre for Computational Design and Discovery of Novel Materials MARVEL, EPFL, Lausanne, Switzerland
⁴IBM Quantum, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA

(Dated: November 4, 2024)

The simulation of molecular electronic structure is an important application of quantum devices. Recently, it has been shown that quantum devices can be effectively combined with classical supercomputing centers in the context of the sample-based quantum diagonalization (SQD) algorithm. This allowed the largest electronic structure quantum simulation to date (77 qubits) and opened near-term devices to practical use cases in chemistry toward the hundred-qubit mark. However, the description of many important physical and chemical properties of those systems, such as photo-absorption/-emission, requires a treatment that goes beyond the ground state alone. In this work, we extend the SQD algorithm to determine low-lying molecular excited states. The extended-SQD method improves over the original SQD method in accuracy, at the cost of an additional computational step. It also improves over quantum subspace expansion based on single and double electronic excitations, a widespread approach to excited states on pre-fault-tolerant quantum devices, in both accuracy and efficiency. We employ the extended SQD method to compute the first singlet (S_1) and triplet (T_1) excited states of the nitrogen molecule with a correlation-consistent basis set, and the ground- and excited-state properties of the [2Fe-2S] cluster.

I. INTRODUCTION

The properties of matter are the results of interactions among its fundamental constituents, namely atoms and molecules. Their behavior is governed by quantum mechanics and is connected to the electronic structure problem, which involves solving the time-independent Schrödinger equation. Determining its lowest-energy solution (i.e. the ground state of electrons at a given nuclear geometry) provides important understandings of the equilibrium properties of chemical compounds, including chemical reactivity, product distributions, and reaction rates.

However, the study of the ground state alone is not exhaustive. Many processes – such as the photochemical phenomena involving absorption/emission of photons and transitions between different many-electron wavefunctions – require the inclusion of excited states to be described correctly [1–4]. Moreover, excited states usually possess symmetries that are different from the ground state ones, and may undergo chemical reactions that would be otherwise forbidden [5, 6].

Modeling these phenomena by solving exactly the electronic Schrödinger equation presents a considerable challenge even for contemporary (super)computers, currently limited to system sizes of 26 electrons in 23 orbitals [7]. As a result, it becomes necessary to employ approximate techniques, each with a specific domain of applicability,

balancing between accuracy and computational cost [8–21].

Quantum computers have emerged as an alternative and complementary platform to simulate ground- and excited-state properties of chemical systems. The linear scaling between number of qubits and number of orbitals, combined with the polynomial cost of simulating the time-dependent Schrödinger equation, offers new algorithms (e.g. phase estimation, adiabatic state preparation, and quantum filter diagonalization) to simulate electronic systems [22–28]. While these algorithms have favorable asymptotic scaling, their performance and cost are also affected by constant prefactors, determining the size and depth of the corresponding quantum circuits as well as their accuracy.

As a result, experimental demonstration on near-term quantum processing units (QPUs) have focused on heuristic methods, e.g. based on the variational principle [29–33], that may deliver accurate results in certain regimes and account for hardware compatibility in their design. In this context, usually a QPU executes a subroutine in a larger program managed by a classical computer. Although these approaches have been applied to the simulation of chemical and physical systems, calculations have remained confined to \sim 2-16 orbitals.

These limitations originate in part from the substantial overhead of quantum measurements affecting QPUs [34–36] and to the sensitivity of these measurements to errors, which prompted to reconsider the interaction between classical and quantum devices through the quantum-centric (super)computing model [37]. Pre-, peri- and post–processing by classical computers allows to execute a limited number of large quantum circuits on QPUs, and to mitigate errors at the level of individual sam-

^{*} stefano.barison@epfl.ch

[†] j.robledomoreno@ibm.com

[†] mario.motta@ibm.com

ples from quantum measurements, giving access to large and more challenging instances of the electronic structure problem. For example, the Sample-based Quantum Diagonalization algorithm (SQD) [38–40] uses a QPU to sample electronic configurations and a classical computer to recover signal from such noisy data and to diagonalize the Hamiltonian in a subspace defined by the recovered configurations [38]. SQD allowed ground-state simulations in active spaces of up to 36 orbitals, using up to 77 qubits and 10570 quantum gates [39]. However, its application to electronic excited states is yet to be established.

Here, we present an extension of the SQD algorithm (Ext-SQD) to compute excited states of electronic systems. First, we compare the method to other approaches, assessing its accuracy and efficiency. Then, we use Ext-SQD to simulate the ground and low-lying excited states of N₂ along breaking of the triple bond in a correlationconsistent cc-pVDZ basis set. The experiment is performed using 58 qubits on a Heron quantum processor and 5204 (1792 2-qubit) quantum gates. Finally, we apply the Extended-SQD to the study of the active space electronic structure of a methyl-capped [2Fe-2S] cluster [41, 42] using a (30e,20o) active space, and was simulated using 45 qubits on a Heron quantum processor and 3170 quantum gates (of which 1100 are 2-qubit gates). The structure of this work is as follows: In Section II we introduce different approaches to excited states determination and present our extension to the SQD algorithm. Then, in Section III we present the results of our classical simulations and QPU experiments. Finally, we conclude in Section IV with our considerations and outlooks on the proposed algorithm.

II. METHODS

We start from the Born-Oppenheimer Hamiltonian in second quantization,

$$\hat{H} = E_0 + \sum_{\substack{pr\\ \tau}} h_{pr} \hat{a}_{p\sigma}^{\dagger} \hat{a}_{r\sigma} + \sum_{\substack{pr\\ \tau \neq s}} \frac{(pr|qs)}{2} \hat{a}_{p\sigma}^{\dagger} \hat{a}_{q\tau}^{\dagger} \hat{a}_{s\tau} \hat{a}_{r\sigma} , \quad (1)$$

where p, r, q, s label spatial orbitals φ_p in a finite orthonormal basis of M elements, $\sigma, \tau \in \{\alpha, \beta\}$ label spin polarizations, E_0 is an energy offset, and h_{pr} and (pr|qs) are the one- and two-electron integrals over the orbitals φ_p . Hartree units are used throughout, unless otherwise specified, and the numbers of spin- up and spin-down electrons are N_{α} and N_{β} , respectively. The goal of this study is to solve the time-independent Schrödinger equation for the Hamiltonian Eq. (1), $\hat{H} |\Phi_{\mu}\rangle = \varepsilon_{\mu} |\Phi_{\mu}\rangle$, to determine its ground and low-lying excited states.

Here, we map the fermionic degrees of freedom of Eq. (1) to qubits with a Jordan-Wigner (JW) transformation [43–45]. In the JW mapping, the single-qubit basis states $|0\rangle/|1\rangle$ represent empty/occupied spin-orbitals, and therefore a computational basis state, i.e. $|\mathbf{x}\rangle$ with

 $\mathbf{x} \in \{0,1\}^{2M}$, represents a Slater determinant of the form

$$|\mathbf{x}\rangle = \prod_{p\sigma} \left(\hat{a}_{p\sigma}^{\dagger}\right)^{x_{p\sigma}} |\varnothing\rangle ,$$
 (2)

where $|\varnothing\rangle$ is the vacuum state. The wavefunction Eq. (2) has N_{σ} spin- σ electrons if $N_{\sigma}(\mathbf{x}) = \sum_{p} x_{p\sigma} = N_{\sigma}$, i.e., the first and second half of the bitstring \mathbf{x} have Hamming weights N_{α} and N_{β} , respectively. Therefore, the target Hamiltonian eigenstates are linear combinations $|\Phi_{\mu}\rangle = \sum_{\mathbf{x}} c_{\mathbf{x}\mu} |\mathbf{x}\rangle$ of computational basis states with such properties. In this work, we will call the states in Eq. (2) "configurations", although in general they are not configuration state functions, i.e., eigenstates of the total spin operator.

A. Excited-state methods

SQD – The first method used in this work to study Hamiltonian excited states is the Sample-based Quantum Diagonalization (SQD) [39, 46] itself. Within SQD, a quantum circuit Ψ is used to sample a set $\tilde{\mathcal{X}} = \{\mathbf{x}_1' \dots \mathbf{x}_N'\}$ of computational basis states. On a pre-fault-tolerant device, even if the the quantum circuit Ψ does not break the particle-number and spin-z symmetries of Eq. (1), due to decoherence, $\tilde{\mathcal{X}}$ may contain configurations with incorrect particle numbers, $N_{\sigma}(\mathbf{x}_k') \neq N_{\sigma}$ for some $\mathbf{x}_k \in \tilde{\mathcal{X}}$. To overcome this limitation, Ref. [39] proposed a self-consistent configuration recovery (S-CORE), that allows to transform $\tilde{\mathcal{X}}$ into a self-consistently determined set of configurations $\mathcal{X}_{\mathrm{R}} = \{\mathbf{x}_1 \dots \mathbf{x}_N\}$ with correct particle number, i.e. $N_{\sigma}(\mathbf{x}_k) = N_{\sigma}$ for all $\mathbf{x}_k \in \mathcal{X}_{\mathrm{R}}$.

To approximate the ground-state wave function, K sets of configurations or "batches" are randomly sampled from $\mathcal{X}_{\mathbf{R}}$, and each batch is used to produce a set $S^{(b)} = \{\mathbf{y}_1^{(b)} \dots \mathbf{y}_D^{(b)}\}$ of configurations closed under spin inversion symmetry [39]. Note that, for this reason, D can be substantially larger than $|\tilde{\mathcal{X}}|$. Within each batch, the ground-state wave function is approximated as

$$|\tilde{\Phi}_{0}^{(b)}\rangle = \sum_{k=1}^{D} c_{k0}^{(b)} |\mathbf{y}_{k}^{(b)}\rangle ,$$
 (3)

where the coefficients $c_{k0}^{(b)}$ are determined by solving the eigenvalue equation

$$\sum_{k=1}^{D} H_{kl}^{(b)} c_{k0}^{(b)} = \tilde{\varepsilon}_{0}^{(b)} c_{l0}^{(b)} , \ H_{kl}^{(b)} = \langle \mathbf{y}_{l}^{(b)} | \hat{H} | \mathbf{y}_{k}^{(b)} \rangle , \qquad (4)$$

and $\tilde{\varepsilon}_0^{(b)}$ is the lowest eigenvalue of $H^{(b)}$. The wave function Eq. (3) leading to the lowest energy $\tilde{\varepsilon}_0^{(b)}$ is chosen as ground-state approximation. We will henceforth denote the corresponding set of configurations, eigenvalue, and eigenvector as $S = \{\mathbf{y}_1 \dots \mathbf{y}_D\}$, $|\tilde{\Phi}_0\rangle$, and $\tilde{\varepsilon}_0$ respectively.

A simple and relatively economical way to approximate low-lying excited states of Eq. (1) within SQD is to compute multiple eigenpairs of the matrix H, which yields wave functions of the form

$$|\tilde{\Phi}_{\mu}\rangle = \sum_{k=1}^{D} c_{k\mu} |\mathbf{y}_{k}\rangle .$$
 (5)

This procedure, though natural and compelling, has an important limitation: configurations in $\tilde{\mathcal{X}}$ are drawn from a quantum circuit Ψ that approximates the ground state of \hat{H} , and configurations \mathcal{X}_R are derived from $\tilde{\mathcal{X}}$ through a S-CORE procedure guided by approximate ground-state wavefunctions. These properties may strongly bias SQD towards the ground-state wave function leading, e.g., to overestimate excitation energies. Furthermore, when a Hamiltonian has a molecular point-group symmetry (e.g. C_{2v} or $D_{\infty h}$), the set \mathcal{X}_R contains configurations in a specific irrep of the symmetry group (e.g. A_1 or A_{1g} respectively) and excited states in different irreps are not accessible within SQD.

QSE(SD) – An alternative to SQD is a quantum subspace expansion (QSE) [28, 47] based on single and double electronic excitations. In this framework, Hamiltonian eigenstates are approximated by states of the form

$$|\tilde{\Phi}_{\mu}\rangle = \sum_{I} d_{I\mu} \,\hat{E}_{I} \,|\tilde{\Phi}_{0}\rangle \tag{6}$$

where $\hat{E}_I \in \{\mathbb{I}, \hat{a}_{a\sigma}^{\dagger} \hat{a}_{i\sigma}, \hat{a}_{a\sigma}^{\dagger} \hat{a}_{b\tau}^{\dagger} \hat{a}_{j\tau} \hat{a}_{i\sigma}\}$ is a set of excitation operators – here, single and double excitations from occupied to virtual orbitals, respectively ij and ab, defined by the restricted closed-shell Hartree-Fock state – and $|\tilde{\Phi}_0\rangle$ is the SQD ground-state approximation Eq. (3). The coefficients $d_{I\mu}$ are determined variationally, solving the generalized eigenvalue equation $Md = Sd\tilde{\epsilon}$ with

$$M_{IJ} = \langle \tilde{\Phi}_0 | \hat{E}_I^{\dagger} \hat{H} \hat{E}_J | \tilde{\Phi}_0 \rangle , S_{IJ} = \langle \tilde{\Phi}_0 | \hat{E}_I^{\dagger} \hat{E}_J | \tilde{\Phi}_0 \rangle .$$
 (7)

This method, here abbreviated as QSE(SD), has been used to simulate molecular excited states on quantum hardware in the last few years [48–55], in part because it does not require running deeper circuits than those needed to prepare $|\tilde{\Phi}_0\rangle$. However, QSE(SD) has a substantial measurement overhead, as it requires estimating high-order k-body reduced density matrices. Furthermore, it shares the accuracy limitations of truncated configuration interaction methods, e.g. the lack of size-extensitivity and size-consistency of ground-state energies and lack of size-intensivity of excitation energies [28].

Ext-SQD – A third path toward to compute molecular excited states is an extension of the SQD method, rather naturally suggested by the QSE(SD) approach. The approximate eigenfunctions in Eq. (6) have the following form, readily obtained from Eq. (3),

$$|\tilde{\Phi}_{\mu}\rangle = \sum_{Ik} d_{I\mu} c_{k0} \,\hat{E}_I \,|\mathbf{y}_k\rangle \ .$$
 (8)

Since the action of an excitation operator \hat{E}_I on a computational basis state $|\mathbf{y}_k\rangle$ yields another computational basis state,

$$\hat{E}_I | \mathbf{y}_k \rangle = \gamma_{Ik} | \mathbf{z}_{Ik} \rangle , \ \gamma_{Ik} \in \{0, 1, -1\} ,$$
 (9)

equation (8) takes the form

$$|\tilde{\Phi}_{\mu}\rangle = \sum_{Ik} d_{I\mu} \gamma_{Ik} c_{k0} |\mathbf{z}_{Ik}\rangle .$$
 (10)

Equation (10) shows that $|\tilde{\Phi}_{\mu}\rangle$ is a linear combination of computational basis state in the extended set $S_{\rm E} = \{\mathbf{z}_{Ik}\}_{Ik}$, with particularly structured expansion coefficients $d_{I\mu}\gamma_{Ik}c_{k0}$. In the light of Eq. (10), it is natural to consider approximations to excited states of the form

$$|\tilde{\Phi}_{\mu}\rangle = \sum_{\mathbf{z} \in S_{\mathrm{E}}} f_{\mathbf{z}\mu} |\mathbf{z}\rangle ,$$
 (11)

i.e., linear combinations of computational basis state in $S_{\rm E}$ with coefficients determined variationally exactly as in SQD, Eq. (4). We call this approach Extended SQD (Ext-SQD), and employ it along SQD and QSE(SD) in the reminder of this work.

Compared to SQD, Ext-SQD is by construction more accurate and expensive, because it uses configurations in $S_{\rm E}$ rather than $\mathcal{X}_{\rm R}$. The increase in computational cost is relatively modest, since $D_E = |S_E|$ is bounded by $D \prod_{\sigma} N_{\sigma}(M - N_{\sigma})$. Compared to QSE(SD), Ext-SQD is by construction more accurate because, while both use computational basis states in $S_{\rm E}$, Ext-SQD has expansion coefficients, Eq. (11), free from the particular form in Eq. (10), which is a source of greater variational flexibility. Furthermore, Ext-SQD is less expensive than QSE(SD), because it does not require the computation of high-order k-body reduced density matrices. It is also less prone to numeric instability, because it requires solving a standard eigenvalue equation rather than a generalized eigenvalue equation, Eq. (7), where the overlap matrix in Eq. (7) may be ill-conditioned.

B. Computational details

1. Active spaces

For the dissociation of N_2 , we used the classical electronic structure PySCF [56, 57] to generate optimized mean-field orbitals (of restricted closed-shell Hartree-Fock, RHF, type) and matrix elements of the Hamiltonian Eq. (1) in:

1. a (10e,80) active space obtained from the cc-pVDZ basis set [58] using the automated valence active-space (AVAS) procedure [59]; we selected active-space orbitals overlapping with nitrogen 2s and 2p atomic orbitals

- a (10e,16o) active space obtained from the 6-31G basis set [60] using the frozen-core approximation (results from this active space are presented exclusively in the Appendix)
- 3. a (10e,26o) active space obtained from the cc-pVDZ basis set using the frozen-core approximation

For the [2Fe-2S] cluster, we used a publicly available (30e,20o) active space [61] comprising iron 3d, sulfur 3p, and cluster-ligand bonding orbitals [41, 42].

The SQD and Ext-SQD calculations are carried out using the Qiskit-addon-sqd [62] package. To provide a reference against which to interpret SQD, QSE(SD), and Ext-SQD results, we computed ground- and excited-state energies with the Complete Active-Space Configuration Interaction (CASCI) and Heat-Bath Configuration Interaction (HCI) [63] as implemented in PySCF.

2. Quantum circuits

Having selected a set of single-electron orbitals for each of the studied species, we constructed quantum circuits from which to measure configurations. Following [39], in this work we employed a truncated version of the local unitary cluster Jastrow (LUCJ) ansatz [64],

$$|\Psi\rangle = e^{-\hat{K}_1} e^{\hat{K}_0} e^{i\hat{J}_0} e^{-\hat{K}_0} |\mathbf{x}_{\text{RHF}}\rangle ,$$
 (12)

where \mathbf{x}_{RHF} is the bitstring representing the RHF state in the JW mapping and

$$\hat{K}_{\mu} = \sum_{\substack{pr \\ \sigma}} K_{pr}^{\mu} \, \hat{a}_{p\sigma}^{\dagger} \hat{a}_{r\sigma} \, , \, \, \hat{J}_{\mu} = \sum_{\substack{pr \\ \sigma\sigma}} J_{p\sigma,r\tau}^{\mu} \, \hat{n}_{p\sigma} \hat{n}_{r\tau} \, , \quad (13)$$

are a generic one-body operator and a density-density operator, respectively. The locality approximation underlying LUCJ involves restricting the density-density operators, i.e., sparsifying the matrices J^{μ} to include only spin-orbitals mapped to qubits that are adjacent in the topology of a given device [64]. Alternatively, one could include spin-orbitals mapped to qubits that can be coupled using SWAP networks compatible with the device's topology. Following [39] we performed optimization-free experiments, leveraging the connection between LUCJ and classical coupled-cluster theory (CCSD) to parametrize the circuits in Eq. (12) using t_2 amplitudes from a CCSD calculation, however performing a quantum-classical optimization could improve the quality of the results.

The LUCJ circuits were generated using the ffsim library [65] and executed on IBM's 133-qubit Heron superconducting quantum processor ibm_torino, using twirled readout error mitigation (ROEM) [66] and dynamical decoupling (DD) [67–70] to mitigate errors arising from qubit measurements and quantum gates, respectively. We employed the implementation of ROEM and DD available through the SamplerV2 primitive of Qiskit's [71] Runtime library. See Appendix B for additional details.

3. Details of SQD, QSE(SD), and Ext-SQD calculations

After executing the quantum circuit Eq. (12), we run SQD to approximate the molecular ground state. The number of configurations gathered, iterations of S-CORE, number of batches, dimension of the SQD subspace, Ext-SQD, and number of single and double excitations (specifying the cost of QSE(SD)) for N₂ are listed in Table I.

system	$ \tilde{\mathcal{X}} $	S-CORE	K	D	D_E	(n_s, n_d)
N_2 (10e,8o)	10^{3}	5	10	$2.5\cdot 10^3$	$3 \cdot 10^3$	(56,210)
N_2 (10e,16o)	10^{5}	10	10	$4 \cdot 10^6$	$2.1\cdot 10^6$	(272,5460)
N_2 (10e,26o)	10^{5}	10	10	$9 \cdot 10^6$	$4.7\cdot 10^6$	(702,44850)

TABLE I. Summary of the N_2 Ext-SQD experiments. The table shows the different active spaces employed in the SQD calculations of N_2 . $|\tilde{\mathcal{X}}|$ indicates the amount of samples (either classical or quantum) collected, while S-CORE and K the number of configuration recovery iterations run and the number of different batches used, respectively. D indicates the final dimension of the subspace in which the classical solver is run. In case of multiple instances, the maximum D is shown. Finally, (n_s, n_d) indicates the number of single and double excitations operators used in the excited-state calculations, while D_E the corresponding extended subspace dimension. While extended, D_E might be smaller than D in some instances (see Appendix A for additional information).

We use the classical solver to determine not only the ground state, but also the eigenpairs of the first two excited states, T_1 and S_1 . T_1 is the first triplet state, while S_1 is the first singlet excited state. QSE(SD) is run with all single and double excitation operators, forming square matrices M_{IJ} and S_{IJ} of Eq. (7) of dimension $(n_s+n_d+1)\times(n_s+n_d+1)$. Finally, Ext-SQD is performed by applying all single and double excitation operators to the SQD wave function.

For the [2Fe-2S] cluster, $|\tilde{\mathcal{X}}| = 2.4576 \cdot 10^6$, S-CORE = 7, $D = 16.7 \cdot 10^6$, $D_E = 126.2 \cdot 10^6$ and $(n_s, n_d, n_t) = (420, 14535, 610470)$, where n_t is the number of three-particle excitations.

III. RESULTS

A. Assessment of accuracy

As highlighted in Section II A, there are several approaches to compute excited states based on SQD. We compare them in Figure 1, focusing on the dissociation of N_2 molecule in the (10e,8o) active space described in Sec. II B 1. The calculations in Figure 1 use a classical device to draw configurations from the uniform probability distribution, and perform SQD calculations as detailed in Sec. II B 3. We use QSE(SD), Ext-SQD, and CASCI to determine the ground state (GS) and the first excited state (T₁) of the molecule. The latter is the lowest-energy

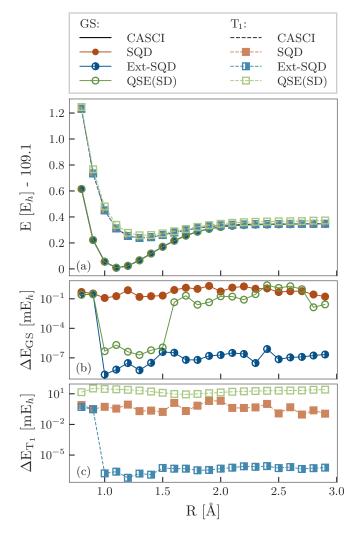


FIG. 1. **Ext-SQD** assessment of accuracy. (a) Ground-state (GS) and first excited-state (T₁) potential energy surfaces of N₂ in a (10e,8o) active space using the SQD method (SQD), its extension (Ext-SQD), the quantum subspace expansion with single and double excitation operators (QSE(SD)), and classical diagonalization (CASCI). (b,c) Deviations ΔE from CASCI in the total energy, for SQD, Ext-SQD, and QSE(SD), along the GS (b) and T₁ (c) potential energy curves.

triplet state. As highlighted in the middle and bottom panels, both QSE(SD) and Ext-SQD deviate from FCI by less than 1 mHa across dissociation. On the other hand, QSE(SD) overestimates the energy of T_1 by 10-30 mHa whereas Ext-SQD agrees with CASCI within 1 mHa.

While QSE(SD) requires solving an eigenvalue equation of size 267, smaller than Ext-SQD, the computational cost of forming the Hamiltonian and overlap matrix is higher than that of producing the Ext-SQD wavefunction. Indeed, every point of the dissociation curve for QSE(SD) required 5 hours running on 30 cores on a bare metal node consisting on four sockets each with a Intel Xeon Platinum 8260 (2.40GHz) processor, while

the Ext-SQD required on average 20s per point on a single core of an M3 Pro laptop processor. Furthermore, as R increases beyond 2.5 Angstrom, the QSE(SD) overlap matrix is more likely to become ill-conditioned, requiring regularization of the eigenvalue equation. For this reason, QSE(SD) energies are above SQD in that regime. Due to the lower accuracy and higher computational cost of QSE(SD), we henceforth employ SQD and Ext-SQD.

B. Applications: dissociation of N₂

We now consider the dissociation of N_2 in a (10e,26o) active space. The spectrum of N_2 is complex, encompassing electronic transitions over a broad range of wavelengths. It is also thoroughly studied, both experimentally and in classical quantum chemistry [72–76], where is a well-known test of the accuracy of electronic structure methods in the presence of static electronic correlation [77].

The comparison between the SQD, Ext-SQD, and SCI is shown in Fig. 2, focusing on the ground state, T_1 , and S_1 excited states. As seen, Ext-SQD improves the quality of all three curves by order 10^2 milliHartre, and also leads to less pronounced discontinuities as the bondlength R varies. Ext-SQD excited states also feature values of S^2 much closer to the exact values (0 and 2 for singlet and triplet states, respectively) than SQD.

From the results in Fig. 2, we can compute several quantities of interest for nitrogen, and compared it to classical simulations and experimental data. Table II reports the equilibrium bondlengths, vibrational frequencies, and dissociation energies of N₂ obtained with HCI, SQD and Ext-SQD. We estimate equilibrium bondlengths and vibrational frequencies by fitting the potential energy curve in the interval $0.9 \text{ Å} \leq R \leq 1.5 \text{ Å}$ to a Morse potential. We estimate dissociation energies by (1) fitting the potential energy curve to a powerlaw curve in the interval $R < 2.0 \,\text{Å}$, and (2) taking the difference between the asymptotic value of the energy (from the power-law fit) to the value of the energy at the equilibrium bondlength (from the Morse-potential fit). Ext-SQD properties are in substantially better agreement with HCI than their SQD counterparts, illustrating how Ext-SQD can economically improve the SQD ground-state approximation without (or before) resorting to energy-variance extrapolations [39].

C. Applications: the [2Fe-2S] cluster

Given the complicated nature of the electronic-structure properties of the [2Fe-2S] cluster [41], we asses the accuracy of two flavors of Ext-SQD. The first is obtained by the action of all single- and double-particle transition operators (SD), as in Eq. (9), and the second by the action of all single-, double and triple-particle transition operators (SDT).

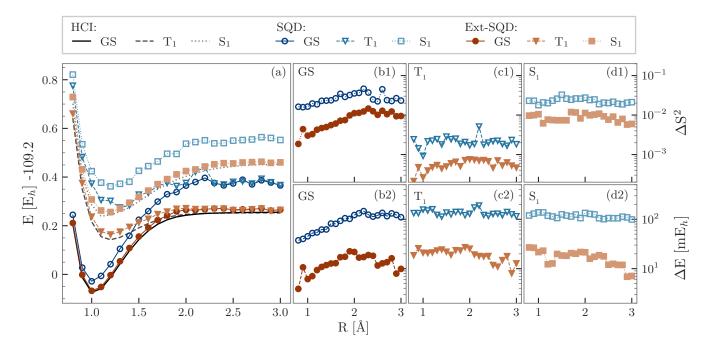


FIG. 2. N_2 dissociation. (a) Ground-state (GS) and first two excited-state potential energy curves (T_1 and S_1) of N_2 in the cc-pVDZ basis set from SQD, extended SQD (Ext-SQD), and heat-bath configuration interaction (HCI). (b1, c1, d1) Deviations ΔS^2 from HCI in the total spin, for SQD and Ext-SQD, along the GS (b1), T_1 (c1), and S_1 (d1) potential energy curves. (b2, c2, d2) Deviations ΔE from HCI in the total energy, for for SQD and Ext-SQD, along the GS (b2), T_1 (c2), and S_1 (d2) potential energy curves.

Quantity	Experiment	HCI	\mathbf{SQD}	Ext-SQD	
Equilibrium bondlength [Å]	1.097513	1.119256	1.109704	1.119951	
Vibrational frequency $[cm^{-1}]$	2330.00	2319.24	2402.53	2319.01	
Dissociation energy [kJ/mol]	945.33 ± 0.59	865.50 ± 0.68	1041.56 ± 7.44	891.87 ± 3.93	

TABLE II. **Properties from the N**₂ **potential energy curve.** Comparison of N₂ equilibrium bondlengths, vibrational frequency, and dissociation energy (rows 2-4) from experiment (column 2, from Refs. [78, 79]) and HCI, SQD, Ext-SQD (columns 3-5) in the cc-pVDZ basis set with frozen-core approximation. Statistical uncertainties in columns 3-5 reflect uncertainties from a standard fitting procedure.

The comparison between HCI, SQD and Ext-SQD ground state energies as a function of the subspace dimension is shown in panel (a) of Fig. 3. For a fixed subspace dimension, the ground-state energy estimations obtained from Ext-SQD (SD) and Ext-SQD (SDT) are in good agreement with those produced by HCI. The passage from SQD to Ext-SQD (SD) decreases the deviation from the ground-state energy – computed with density matrix renormalization group (DMRG) [42] – from ~ 200 to ~ 75 milliHartree, while the number of configurations increases from ~ 15 to ~ 35 million. The inclusion of triple-particle transition operators further decreases the deviation to ~ 50 milliHartree while increasing the number of configurations to ~ 100 million.

Panel (b) of Fig. 3 compares the S_1 and S_2 energies and subspace dimensions obtained from Ext-SQD (SDT) to those obtained with HCI. The points of highest subspace dimension considered in the Ext-SQD (SDT) calculations show excellent agreement (< 1 milliHartree) with the

HCI estimations at comparable subspace dimensions.

In the canonical interpretation of iron-sulfur clusters based on the Heisenberg-Double-Exchange (HDE) model, the Fe atoms are considered in definite valence states. For the oxidized [2Fe-2S]²⁻ cluster considered in this work, both Fe atoms are assumed high-spin Fe(III) with total spins $S_1, S_2 = 5/2$, and it is assumed that there is no double-exchange. Therefore, the HDE model reduces to the Heisenberg form $\hat{H} = 2J\vec{S}_1\cdot\vec{S}_2$ where J>0 is a coupling constant. The ground state of the HDE Hamiltonian is low-spin, with maximal antiferromagnetic alignment.

In Table III, we examine the ability of computational methods, including SQD and Ext-SQD, to describe the ground-state physical properties of the oxidized [2Fe-2S] cluster. Specifically, we compute the number of spin- σ

electrons on Fe atom a = 1, 2 as the expectation value of

$$\hat{N}_{e,\sigma}^{(a)} = \sum_{p \in Fe_a} \hat{n}_{p\sigma} , \qquad (14)$$

where $p \in \text{Fe}_a$ denotes the 3d orbitals of iron atom a, the local spin of Fe atom a as the expectation value of

$$\hat{\mathbf{S}}^{(a)} = \frac{1}{2} \sum_{p \in \mathrm{Fe}_a} \mathbf{P}_{\sigma\tau} \, \hat{a}_{p\sigma}^{\dagger} \hat{a}_{p\tau} \,, \tag{15}$$

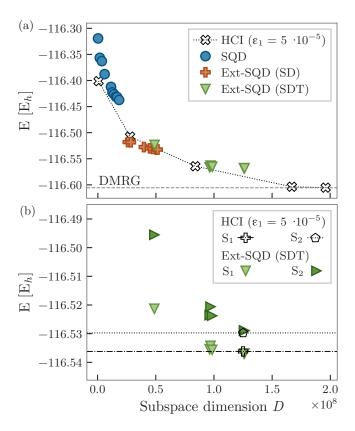
where \mathbf{P} is the vector of 2×2 Pauli matrices, and the spin-spin connected correlation

$$C = \langle \Psi | \hat{\mathbf{S}}^{(1)} \cdot \hat{\mathbf{S}}^{(2)} | \Psi \rangle - \langle \Psi | \hat{\mathbf{S}}^{(1)} | \Psi \rangle \cdot \langle \Psi | \hat{\mathbf{S}}^{(2)} | \Psi \rangle . \quad (16)$$

Restricted HF, MP2, CCSD, and CISD predict \sim 6.2 electrons on each Fe atom, zero local spins, and a very low antiferromagnetic (C < 0) spin-spin correlation. Unrestricted methods predict ~ 5 spin-up and ~ 1 spin-down electrons on the left Fe atom and viceversa on the right Fe atom, local spins of roughly equal magnitude and opposite orientation (roughly ± 2 for the left and right atom, respectively), and a very low antiferromagnetic spin-spin correlation ($C\simeq 0$ because $\langle\Psi|\hat{\mathbf{S}}^{(1)}\cdot\hat{\mathbf{S}}^{(2)}|\Psi\rangle\simeq$ $\langle \Psi | \hat{\mathbf{S}}^{(1)} | \Psi \rangle \cdot \langle \Psi | \hat{\mathbf{S}}^{(2)} | \Psi \rangle$). The behavior of restricted HF, MP2, CCSD, and CISD is a reflection of their inability to capture the strong electronic correlation arising from the composition of the local Fe spins into a global singlet state. Unrestricted methods exhibit spin symmetry breaking with non-zero expectation values for the local spin operators, whose opposite orientation is a reflection of the antiferromagnetic correlation between the two Fe atoms. The values of $\langle \Psi | \hat{\mathbf{S}}^{(1)} \cdot \hat{\mathbf{S}}^{(2)} | \Psi \rangle$ range between -3.84 and -4.41, in good agreement with the DMRG estimate of $\langle \Psi | \hat{\mathbf{S}}^{(1)} \cdot \hat{\mathbf{S}}^{(2)} | \Psi \rangle = -4.92$ [41]. SQD and Ext-SQD predict a stronger antiferromagnetic correlation between the two Fe atoms than other restricted methods, with values of -0.6315 and -1.7619 respectively. Ext-SQD prediction is in agreement with HCI result of -1.7807, which is obtained with a larger subspace dimension. Table IV compares the charge and spin properties of the Fe atoms for the S₁ and S₂ excitations obtained from HCI and Ext-SQD (SDT). As it could be anticipated from the agreement in the energies shown in Fig. 3 (b), the charge and spin observables are also in close agreement. The complete occupation numbers of all the atomic orbitals can be found in the Appendix.

IV. CONCLUSION AND OUTLOOK

In this work, we explored the generalization of sample-based quantum diagonalization (SQD) to the calculation of molecular excited states. We considered three strategies: (i) the calculation of multiple Hamiltonian eigenpairs (beyond the ground-state one) within an SQD calculation, (ii) the quantum subspace expansion obtained applying single and double excitation operators



Ground- and excited-state energies of the FIG. 3. [2Fe-2S] cluster. (a) Ground-state energy estimation as a function of the subspace dimension for SQD (blue circles), Ext-SQD with added single- and double-excitations (SD, orange pluses), and Ext-SQD with single-, double- and tripleexcitations (SDT, green triangles). The energies and subspace dimensions are compared to those visited by different iterations (following the flow of the black dotted line) of HCI (black crosses). The truncation coefficient for the HCI calculation is indicated in the legend (see Ref. [80] for details). The gray horizontal line shows the DMRG energy reported in Ref. [42]. (b) Excited-state energies estimation as a function of the subspace dimension Ext-SQD with single-, double- and triple-excitations (SDT, downward triangles for S_1 , rightward triangles for S_2). Energies and dimensions are compared to HCI results. Dotted and dash-dotted lines represent again the HCI results and are a guide to the eyes. The truncation coefficient for HCI is the same as in panel (a).

(QSE(SD)), and (iii) the extension of SQD where single and double excitation operators are applied to the individual configurations in the SQD wavefunction, yielding a larger set of configurations where used to compute multiple Hamiltonian eigenpairs (Ext-SQD).

Ext-SQD yielded the most accurate excited states, significantly outperforming SQD and QSE(SD) thanks to a greater variational flexibility. Finally, the computational overhead that Ext-SQD introduces over SQD is considerably modest compared to QSE(SD).

Because Ext-SQD returns approximate Hamiltonian eigenstates in the classically accessible form of sparse CI

Method	Fe left			Fe right				
	$N_{e,\uparrow}$	$N_{e,\downarrow}$	Spin	$N_{e,\uparrow}$	$N_{e,\downarrow}$	Spin	Spin correlation	
Restricted								
RHF	3.1075	3.1075	0	3.1071	3.1071	0	-0.0437	
MP2	3.1112	3.1112	0	3.1106	3.1106	0	-0.0183	
CISD	3.1337	3.1337	0	3.1327	3.1327	0	-0.0560	
CCSD	3.1397	3.1397	0	3.1393	3.1393	0	-0.2085	
HCI (D = 196.1 M)	3.0103	3.0103	0	3.0089	3.0089	0	-1.7807	
Unrestricted	Unrestricted							
UHF	4.9970	0.7993	2.0988	0.7966	4.9970	-2.1001	-0.0036	
UMP2	4.9960	0.8285	2.0837	0.8258	4.9960	-2.0851	-0.0011	
UCISD	4.9734	1.0132	1.9801	1.0104	4.9735	-1.9815	-0.0112	
UCCSD	4.9515	1.0347	1.9583	1.0320	4.9516	-1.9598	-0.0288	
Quantum								
$\overline{\mathrm{SQD}\ (D = 16.7\ \mathrm{M})}$	3.1121	3.1121	0	3.1116	3.1116	0	-0.6315	
Ext-SQD $(D = 126.2 \text{ M})$	3.0130	3.0130	0	3.0118	3.0118	0	-1.7619	

TABLE III. **Ground-state charge and spin properties of the [2Fe-2S] cluster.** Average number of spin-up and spin-down electrons and local spin-z on the left (columns 2-4) and right (coulmns 5-7) Fe atom of the [2Fe-2S] cluster, and correlation between local spins on the two Fe atoms (column 8), from restricted (R, rows 4-8) and unrestricted (U, rows 10-3) Hartree-Fock (HF), Møller-Plesset 2nd order perturbation theory (MP2), coupled-clusted singles and doubles (CCSD), and configuration interaction singles and doubles (CISD). Row 8 reports the result for the largest HCI calculation shown in panel (a) of Fig. 3. Rows 14 and 15 correspond to the quantities from SQD and Ext-SQD (SDT), respectively.

Exc.	Fe left			F	e right					
	${ m N}_{e,\uparrow}$	$N_{e,\downarrow}$	Spin	${ m N}_{e,\uparrow}$	$\mathrm{N}_{e,\downarrow}$	Spin	Spin corr.			
HCI (D = 196.1 M)										
S_1	3.1030	3.1030	0	3.0767	3.0767	0	-1.3600			
S_2	3.0966	3.0966	0	3.1021	3.1021	0	-1.1954			
Ext-SQD $(D = 126.2 \text{ M})$										
S_1	3.1219	3.1219	0	3.1216	3.1216	0	-1.0102			
S_2	3.0904	3.0904	0	3.0892	3.0892	0	-1.3646			

TABLE IV. Excited-state charge and spin properties of excited states of the [2Fe-2S] cluster. Columns show the same physical observables as in Table III for the S_1 and S_2 excited states both for HCI and Ext-SQD (SDT).

vectors (i.e. linear combinations of a polynomial number of configurations), users can easily certify and rank the quality of Ext-SQD excited states – e.g., by computing their energy variances and mean energies [81] – and use Ext-SQD wavefunctions as input in further classical post-processing. This includes the computation of various excited-state properties, transition properties (e.g. 1-body reduced density matrices and dipoles), and frequency-dependent response functions (although it should be noted that a large number of excited states, often with energy well above that of T_1 or S_1 , are necessary for accurate computation of these quantities).

The results of this study indicate that Ext-SQD can be used to compute singlet-singlet and singlet-triplet gaps relevant for, e.g., thermally activated delayed fluorescence [82, 83] and photoisomerization [84, 85] – as well as other energy differences like ionization potentials, electron affinities, and band gaps. In addition, due to the possibility of computing excited-state and transition properties, SQD can become a useful tool in the computation of frequency-dependent response functions – including dipole spectra, density-density correlation functions, spin-spin correlation functions, and Green's functions. Importantly, Ext-SQD allows quantum computing studies in active spaces with considerably more electrons and orbitals than previously possible, because they employ classical and quantum computational resources in synergy and remove the need of measuring high-order density matrices, a bottleneck of widespread forms of quantum subspace calculations [28].

Finally, we remark that Ext-SQD can be transported to Hamiltonians that are linear combinations of polynomially many Pauli or k-body fermionic operators (e.g. the Ising, Heisenberg, and Hubbard Hamiltonians) without major algorithmic modifications, and can be used to approximate their low-lying Hamiltonian eigenstates (assuming their sparsity). Therefore, it can be used beyond electronic structure and contribute more broadly to the simulation of many-body quantum systems on quantum computers.

ACKNOWLEDGEMENTS

We acknowledge useful discussion with Antonio Mezzacapo and Kunal Sharma. The code used to produce

- N. A. Anderson, J. J. Shiang, and R. J. Sension, J. Phys. Chem. A 103, 10730 (1999).
- [2] T. Polívka and V. Sundström, Chem. Rev 104, 2021 (2004).
- [3] Y.-C. Cheng and G. R. Fleming, Annu. Rev. Phys. Chem 60, 241 (2009).
- [4] P. J. M. Johnson, A. Halpin, T. Morizumi, V. I. Prokhorenko, O. P. Ernst, and R. J. D. Miller, Nat. Chem 7, 980 (2015).
- [5] R. B. Woodward and R. Hoffmann, J. Am. Chem. Soc 87, 395 (1965).
- [6] T. A. Seder, A. J. Ouderkirk, and E. Weitz, J. Chem. Phys 85, 1977 (1986).
- [7] H. Gao, S. Imamura, A. Kasagi, and E. Yoshida, J. Chem. Theory Comput 20, 1185 (2024).
- [8] E. Runge and E. K. Gross, Phys. Rev. Lett 52, 997 (1984).
- [9] L. Reining, V. Olevano, A. Rubio, and G. Onida, Phys. Rev. Lett 88, 066404 (2002).
- [10] L. Sham and T. Rice, Phys. Rev 144, 708 (1966).
- [11] W. Hanke and L. Sham, Phys. Rev. B 21, 4656 (1980).
- [12] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys 74, 601 (2002).
- [13] K. Emrich, Nuc. Phys. A **351**, 379 (1981).
- [14] J. F. Stanton and R. J. Bartlett, J. Chem. Phys 98, 7029 (1993).
- [15] J. McClain, J. Lischner, T. Watson, D. A. Matthews, E. Ronca, S. G. Louie, T. C. Berkelbach, and G. K.-L. Chan, Phys. Rev. B 93, 235139 (2016).
- [16] J. McClain, Q. Sun, G. K.-L. Chan, and T. C. Berkel-bach, J. Chem. Theory Comput 13, 1209 (2017).
- [17] F. Ma, S. Zhang, and H. Krakauer, New J. Phys 15, 093017 (2013).
- [18] N. Blunt, S. D. Smart, G. H. Booth, and A. Alavi, J. Chem. Phys 143 (2015).
- [19] B. O. Roos, P. R. Taylor, and P. E. Sigbahn, Chem. Phys 48, 157 (1980).
- [20] B. O. Roos, Int. J. Quantum Chem 18, 175 (1980).
- [21] M. W. Schmidt and M. S. Gordon, Annu. Rev. Phys. Chem 49, 233 (1998).
- [22] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science 309, 1704–1707 (2005).
- [23] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Proc. Nat. Acad. Sci. USA 115, 9456 (2018).
- [24] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, "Quantum computation by adiabatic evolution," (2000), arXiv:quant-ph/0001106 [quant-ph].
- [25] R. Babbush, P. J. Love, and A. Aspuru-Guzik, Sci. Rep 4, 6603 (2014).
- [26] T. Albash and D. A. Lidar, Rev. Mod. Phys. 90, 015002 (2018).
- [27] W. Kirby, M. Motta, and A. Mezzacapo, Quantum 7, 1018 (2023).
- [28] M. Motta, W. Kirby, I. Liepuoniute, K. J. Sung, J. Cohn, A. Mezzacapo, K. Klymko, N. Nguyen, N. Yoshioka, and J. E. Rice, Electron. Struct 6, 013001 (2024).
- [29] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien,

- Nat. Commun 5 (2014).
- [30] N. C. Rubin, "A hybrid classical/quantum approach for large-scale studies of quantum systems with density matrix embedding theory," (2016), arXiv:1610.06910 [quant-ph].
- [31] M. Motta, C. Sun, A. T. K. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Nat. Phys 16, 205–210 (2019).
- [32] A. Eddins, M. Motta, T. P. Gujarati, S. Bravyi, A. Mezzacapo, C. Hadfield, and S. Sheldon, PRX Quantum 3 (2022).
- [33] S.-X. Zhang, Z.-Q. Wan, C.-K. Lee, C.-Y. Hsieh, S. Zhang, and H. Yao, Phys. Rev. Lett. 128, 120502 (2022).
- [34] D. Wecker, M. B. Hastings, and M. Troyer, Phys. Rev. A 92, 042303 (2015).
- [35] J. F. Gonthier, M. D. Radin, C. Buda, E. J. Doskocil, C. M. Abuan, and J. Romero, Phys. Rev. Research 4, 033154 (2022).
- [36] G. Mazzola, J. Chem. Phys **160**, 010901 (2024).
- [37] Y. Alexeev, M. Amsler, M. A. Barroca, S. Bassini, T. Battelle, D. Camps, D. Casanova, Y. J. Choi, F. T. Chong, C. Chung, et al., Future Gener. Comput. Syst 160, 666 (2024).
- [38] K. Kanno, M. Kohda, R. Imai, S. Koh, K. Mitarai, W. Mizukami, and Y. O. Nakagawa, "Quantum-selected configuration interaction: classical diagonalization of hamiltonians in subspaces selected by quantum computers," (2023), arXiv:2302.11320 [quant-ph].
- [39] J. Robledo-Moreno, M. Motta, H. Haas, A. Javadi-Abhari, P. Jurcevic, W. Kirby, S. Martiel, K. Sharma, S. Sharma, T. Shirakawa, I. Sitdikov, R.-Y. Sun, K. J. Sung, M. Takita, M. C. Tran, S. Yunoki, and A. Mezzacapo, "Chemistry beyond exact solutions on a quantum-centric supercomputer," (2024), arXiv:2405.05068.
- [40] D. Kaliakin, A. Shajan, J. R. Moreno, Z. Li, A. Mitra, M. Motta, C. Johnson, A. A. Saki, S. Das, I. Sitdikov, A. Mezzacapo, and K. M. M. J. au2, "Accurate quantum-centric simulations of supramolecular interactions," (2024), arXiv:2410.09209 [quant-ph].
- [41] S. Sharma, K. Sivalingam, F. Neese, and G. K.-L. Chan, Nat. Chem 6, 927 (2014).
- [42] Z. Li and G. K.-L. Chan, J. Chem. Theory Comput 13, 2681 (2017).
- [43] G. Ortiz, J. Gubernatis, E. Knill, and R. Laflamme, Comp. Phys. Comm 146, 302 (2002).
- [44] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A 65, 042323 (2002).
- [45] R. D. Somma, Quantum computation, complexity, and many-body physics, Ph.D. thesis, Instituto Balseiro, S.C. de Bariloche, Argentina and Los Alamos National Labo- ratory, Los Alamos, USA (2005), arXiv:quantph/0512209.
- [46] K. Kanno, M. Kohda, R. Imai, S. Koh, K. Mitarai, W. Mizukami, and Y. O. Nakagawa, arXiv:2302.11320 (2023).
- [47] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. De Jong, Phys. Rev. A 95, 042308 (2017).

- [48] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. Kimchi-Schwartz, J. McClean, J. Carter, W. De Jong, and I. Siddiqi, Phys. Rev. X 8, 011021 (2018).
- [49] B. Huang, M. Govoni, and G. Galli, PRX Quantum 3, 010339 (2022).
- [50] I. T. Khan, M. Tudorovskaya, J. J. M. Kirsopp, D. Muñoz Ramo, P. Warrier, D. K. Papanastasiou, and R. Singh, J. Chem. Phys 158, 214114 (2023).
- [51] D. Dhawan, D. Zgid, and M. Motta, J. Chem. Theory Comput 20, 4629 (2024).
- [52] A. Tammaro, D. E. Galli, J. E. Rice, and M. Motta, J. Phys. Chem. A 127, 817 (2023).
- [53] B. Huang, N. Sheng, M. Govoni, and G. Galli, J. Chem. Theory Comput 19, 1487 (2023).
- [54] M. Motta, G. O. Jones, J. E. Rice, T. P. Gu-jarati, R. Sakuma, I. Liepuoniute, J. M. Garcia, and Y. Ohnishi, Chem. Sci 14, 2915 (2023).
- [55] M. A. Castellanos, M. Motta, and J. E. Rice, Mol. Phys e2282736 (2023).
- [56] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, et al., WIREs Comput. Mol. Sci 8, e1340 (2018).
- [57] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, et al., J. Chem. Phys 153, 024109 (2020).
- [58] T. H. Dunning Jr, J. Chem. Phys **90**, 1007 (1989).
- [59] E. R. Sayfutyarova, Q. Sun, G. K.-L. Chan, and G. Knizia, J. Chem. Theory Comput 13, 4063 (2017).
- [60] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys 56, 2257 (1972).
- [61] Z. Li and G. K.-L. Chan, "Github repository: Active space model for iron sulfur clusters," (2017).
- [62] A. A. Saki, S. Barison, B. Fuller, J. R. Garrison, J. R. Glick, C. Johnson, A. Mezzacapo, J. Robledo-Moreno, M. Rossmannek, P. Schweigert, I. Sitdikov, and K. J. Sung, "Qiskit addon: samplebased quantum diagonalization," https://github.com/ Qiskit/qiskit-addon-sqd (2024).
- [63] A. A. Holmes, N. M. Tubman, and C. Umrigar, J. Chem. Theory Comput 12, 3674 (2016).
- [64] M. Motta, K. J. Sung, K. B. Whaley, M. Head-Gordon, and J. Shee, Chem. Sci 14, 11213 (2023).
- [65] The ffsim developers, "ffsim: faster simulations of fermionic quantum circuits," (2024).
- [66] P. D. Nation, H. Kang, N. Sundaresan, and J. M. Gambetta, Phys. Rev. X Quantum 2, 040326 (2021).
- [67] L. Viola and S. Lloyd, Phys. Rev. A 58, 2733 (1998).
- [68] A. Kofman and G. Kurizki, Phys. Rev. Lett 87, 270405 (2001).
- [69] M. J. Biercuk, H. Uys, A. P. VanDevender, N. Shiga, W. M. Itano, and J. J. Bollinger, Nature 458, 996 (2009).
- [70] S. Niu and A. Todri-Sanial, IEEE Trans. Quantum Eng 3, 1 (2022).
- [71] Qiskit contributors, "Qiskit: An open-source framework for quantum computing," (2023).
- [72] A. Lofthus and P. H. Krupenie, J. Phys. Chem. Ref. Data 6, 113 (1977).
- [73] H. Larsen, J. Olsen, P. Jørgensen, and O. Christiansen, J. Chem. Phys 113, 6677 (2000).
- [74] R. Ono, C. Tobaru, Y. Teramoto, and T. Oda, Plasma Sources Sci. Technol 18, 025006 (2009).
- [75] R. J. Gdanitz, Chem. Phys. Lett 283, 253 (1998).

- [76] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys 116, 4462 (2002).
- [77] I. W. Bulik, T. M. Henderson, and G. E. Scuseria, J. Chem. Theory Comput 11, 3171 (2015).
- [78] K. K. Irikura, J. Phys. Chem. Ref. Data 36, 389 (2007).
- [79] K. Huber, Molecular spectra and molecular structure: IV. Constants of diatomic molecules (Springer Science & Business Media, 2013).
- [80] J. E. T. Smith, B. Mussard, A. A. Holmes, and S. Sharma, J. Theor. Chem. 13, 5468 (2017).
- [81] D. Wu, R. Rossi, F. Vicentini, N. Astrakhantsev, F. Becca, X. Cao, J. Carrasquilla, F. Ferrari, A. Georges, M. Hibat-Allah, et al., arXiv:2302.04919 (2023).
- [82] H. Uoyama, K. Goushi, K. Shizu, H. Nomura, and C. Adachi, Nature 492, 234 (2012).
- [83] Q. Gao, G. O. Jones, M. Motta, M. Sugawara, H. C. Watanabe, T. Kobayashi, E. Watanabe, Y.-y. Ohnishi, H. Nakamura, and N. Yamamoto, npj Comput. Mater 7, 70 (2021).
- [84] A. Bylina and Z. Grabowski, J. Chem. Soc., Faraday Trans 65, 458 (1969).
- [85] J. Saltiel, D. Chang, E. Megarity, A. Rousseau, P. Shannon, B. Thomas, and A. Uriarte, Pure Appl. Chem 41, 559 (1975).

Appendix A: Details of extended-SQD algorithm

This Appendix provides a detailed description of the extended-SQD procedure.

1. Extending the SQD subspace

As explained in Section II A the extended subspace of Ext-SQD is obtained by applying a set of excitation operators \hat{E}_I to the configurations of the SQD wave function. As we consider product of creation and annihilation operators, each excitation operator \hat{E}_I on a computational basis state $|\mathbf{y}_k\rangle$ yields another computational basis state

$$\hat{E}_I | \mathbf{y}_k \rangle = \gamma_{Ik} | \mathbf{z}_{Ik} \rangle , \ \gamma_{Ik} \in \{0, 1, -1\} . \tag{A1}$$

The application of \hat{E}_I to the configuration $|\mathbf{y}_k\rangle$ has four different outcomes:

- 1. the configuration \mathbf{z}_{Ik} is new to the subspace and unique
- 2. the configuration \mathbf{z}_{Ik} is not physical
- 3. the configuration \mathbf{z}_{Ik} is new to the subspace but equal to one already obtained \mathbf{z}_{Jl}
- 4. the configuration \mathbf{z}_{Ik} was already in the SQD subspace.

While the first outcome is the one we are looking for and expected to provide the accuracy improvements we show in the main text, the other three have to be considered. Given that we choose \hat{E}_I to be particle-preserving, the outcome 2. is easily handled by checking that $\gamma_{Ik} = 0$. Outcomes 3. and 4. can be controlled by checking the already existing configurations. This check can be performed after every new configuration \mathbf{z}_{Ik} is produced, which results in time consuming strictly non-parallel operations. This check can also be executed after all the new configurations \mathbf{z}_{Ik} are produced. In this case, the produce is easily parallelizable, but produces a computational overhead on the memory used.

In order to find a good tradeoff between the two cases, we apply the set of excitation operators on batches of configurations taken from the SQD wave function. The size and the number of batches can be tuned depending on the system under study. On a single batch, the application of the excitation operators is executed in parallel. Then, the check is performed inter- and intra-batches serially.

Given a suitable amount of time, the application of excitation operators to produce the Ext-SQD can therefeore be performed on every classical device which is able to store the initial SQD wave function. Finally, as indicated in Section II A, the dimension of the extended subspace D_E is is bounded by $D \prod_{\sigma} N_{\sigma}(M - N_{\sigma})$. Therefore, no asimptotically-hard procedure is introduced by the Ext-SQD workflow.

2. Cutting the SQD subspace

The SQD algorithm produces a classic vector representing the wave function of the physical system in the quantum-selected subspace of the Hilbert space. As shown in [39], SQD is able to produce wave function in very good qualitative agreement with the best classical solvers systems up to 77 qubits. When we look at the accuracies of individual wave function amplitudes produced by SQD, the agreement is exceptional for the larger ones and slowly deteriorates in the description of the tails of the wave function.

While this is to be expected, we can use this behavior to speed up the Ext-SQD calculations without compromising the accuracy of the results. Indeed, before extending the subspace with the excitation operators, one could choose to discard all the configurations that have amplitudes below a certain threshold, effectively reducing the size of the SQD subspace.

The choice of the threshold is dependent on the system under study. For N_2 dissociation in cc-pVDZ (and in 6-31G, see Appendix C) basis set we always employed a threshold of $1 \cdot 10^{-3}$ on the wave function amplitudes for every configuration. The loss in accuracy was always less than a mHa, a decrement well compensated by the final accuracy improvement of Ext-SQD. Panel (a) of Fig. A1 compares the ground-state wave function amplitudes $|c_{0x}|^2$ at R = 2.5 Å obtained with SQD and Ext-SQD, comparing them to CASCI results. As shown in [39], at large bond lengths, with strong static correlation effects, we expect the agreement between SQD and CASCI to be worse than at equilibrium. We observe that below the threshold chosen the accuracy of SQD starts deteriorating, while Ext-SQD amplitudes follow CASCI ones down to 10^{-7} .

Another benefit of the cut of the SQD wave function is illustrated in Table I: the final extended subspace can be smaller than the initial in certain instances. As an example, the largest subspace obtained with Ext-SQD after cutting and extending is roughly a half of the largest SQD one, while the results are always an order of magnitude more accurate. Panel (b) of Fig. A1 provides more details on the results obtained. When the final subspace dimensions of SQD and Ext-SQD are compared, we notice that the final size of Ext-SQD calculations in cc-pVDZ basis set are similar to the size of SQD calculations in 6-31G basis, despite having 10 orbitals more. These results highlight the capability of Ext-SQD to add to the final subspace configuration that are more relevant to ground and low-lying states calculations.

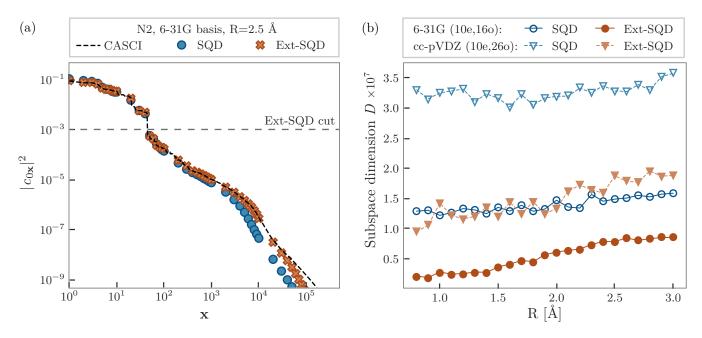


FIG. A1. Cutting in N_2 dissociation. (a) Shows for a selected configuration $R = 2.5 \,\text{Å}$ the GS wave function amplitudes $|c_{0x}|^2$ for the corresponding electronic configuration x. SQD and Ext-SQD are compared to classical CASCI results. The grey dashed line indicates the threshold used to cut the SQD wave function before extending the subspace with single and double excitation operators. For better visualization clarity, the number of configurations shown has been reduced. (b) Shows the dimension of the subspaces obtained with SQD and Ext-SQD in the N_2 hardware experiments along the dissociation path.

Appendix B: Quantum experiment details

This appendix provides details about the experiments conducted on IBM's quantum devices.

The quantum circuits that were considered in these experiments are of the LUCJ type, as defined in Eq. 12 in the main text. Their parameters are obtained from classical CCSD theory, following the procedure used in Ref. [39]. Table A1 provides a summary of the quantum resources involved in the experiments, specifically: the number of qubits and quantum gates and shots $|\tilde{\mathcal{X}}|$ for each circuit, the device used and the specific layout (see Fig. A2 for additional details), the fraction of of the configurations live in the subspace of the Fock space with the correct particle number, defined as:

$$p_N^{\text{hw}} = \frac{1}{|\tilde{\mathcal{X}}|} \sum_{\ell=1}^{|\tilde{\mathcal{X}}|} \delta_{N_{\mathbf{x}_{\ell}\alpha}, N_{\alpha}} \delta_{N_{\mathbf{x}_{\ell}\beta}, N_{\beta}}, \tag{A1}$$

and the fraction of sampled configurations with the correct particle number if samples were collected from the uniform distribution over configurations of length 2M

$$p_N^{\text{unif}} = \binom{M}{N_\alpha} \binom{M}{N_\beta} 2^{-2M}.$$
 (A2)

These two probabilities are statistically distinguishable for all the experiments we carried out.

system	q (JW, Tot.)	(d,CNOT,u)	device	layout	$p_N^{\text{hw}} [95\% \text{ c.i.}]$	$p_N^{ m unif}$	$ ilde{\mathcal{X}} $
N_2 , (10e,16o)	(32, 36)	(148,762,1408)	ibm_nazca	$[0 \rightarrow 82] + [2 \rightarrow 73]$	(0.0069, 0.0071)	0.0044	$1 \cdot 10^5$
N_2 , (10e,26o)	(52, 58)	$(223,\!1792,\!3412)$	ibm_torino	$[0{\to}121]{+}[2{\to}127]$	(0.0016, 0.0017)	$9.6\cdot10^{-7}$	$9.8304\cdot10^4$
[2Fe-2S], (30e,20o)	(40, 45)	(173, 1100, 2070)	ibm_torino	$[0 \rightarrow 104] + [2 \rightarrow 94]$	(0.0044, 0.0045)	0.00022	$2.4576\cdot10^6$

TABLE A1. Hardware details. For the active space of each molecular system (first column), we list the details of the LUCJ quantum circuit executed on hardware: its number of qubits (q) both for the Jordan-Wigner encoding (JW) and the total number of qubits used including auxiliary ones (Tot.) (column 2) and count of quantum operations (circuit depth d, number of CNOT and single-qubit u gates, column 3), the device used (column 4) and the qubit layout chosen (column 5, where $[a \rightarrow b]$ indicates the shortest path between qubits a and b in the device topology (see Fig. A2). In columns 6 and 7 we report a 95% confidence interval for the fraction p_N^{hw} of configurations with the correct particle number and the corresponding value for configurations with uniform probability distribution (see Eqs. A1 and A2). In column 8, we list the number of raw measurement outcomes obtained from the device.

Our experiments used twirled readout error mitigation (ROEM) [66] to mitigate errors arising from qubit measurement. Dynamical decoupling (DD) [67–70] sequences of X control pulses to mitigate errors arising from quantum gates are also employed. We employed the implementation of ROEM and DD available on the Runtime library of Qiskit [71], through the SamplerV2 primitive. The DD sequences consist of two X pulses (as in Ramsey echo experiments) to idle qubits.

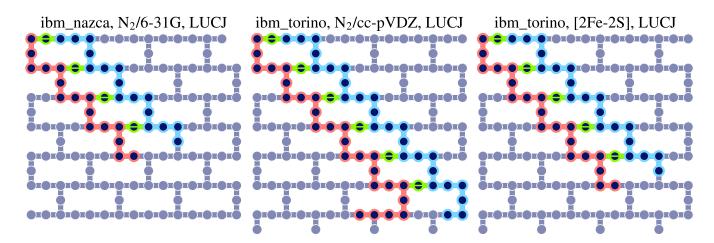


FIG. A2. Experiment layout on IBM quantum processors. Schematics of the devices used to carry out experiments for the N_2 molecule with 6-31G basis, the N_2 molecule with cc-pVDZ basis, and the [2Fe-2S] cluster. Qubits used in the calculation are shown in red (for qubits associated with α spin-orbitals), blue for qubits associated with β spin-orbitals), and green (for auxiliary qubits).

Appendix C: Additional results

This Appendix provides extra experimental results.

1. Dissociation of nitrogen using 6-31G

Hardware experiments for N_2 dissociation were performed in the 6-31G basis set with frozen core approximation. With a total of 10 electrons in 16 orbitals, these experiments represent a intermediate step between the (10e, 8o) and (10e, 26o) active spaces of cc-pVDZ basis set. We collected $|\tilde{\mathcal{X}}| = 10^5$ samples from ibm_nazca, using LUCJ circuits and 36 qubits. The final SQD and Ext-SQD results are shown in Fig. A3, where they are compared against exact classical calculations performed in the same basis set (CASCI).

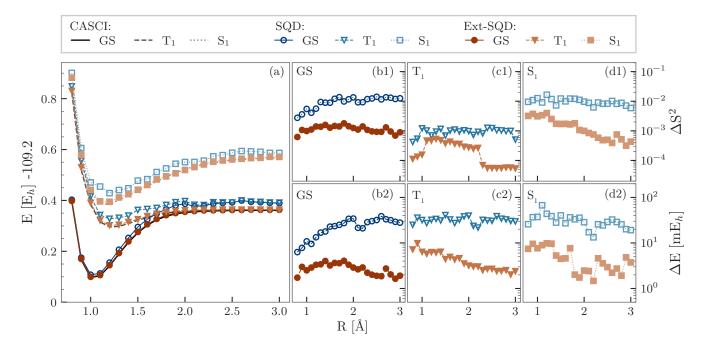


FIG. A3. N_2 dissociation. (a) Ground-state (GS) and first two excited-state potential energy curves (T_1 and S_1) of N_2 in the 6-31G basis set from SQD, extended SQD (Ext-SQD), and heat-bath configuration interaction (HCI). (b1, c1, d1) Deviations ΔS^2 from HCI in the total spin, for SQD and Ext-SQD, along the GS (b1), T_1 (c1), and S_1 (d1) potential energy curves. (b2, c2, d2) Deviations ΔE from HCI in the total energy, for for SQD and Ext-SQD, along the GS (b2), T_1 (c2), and S_1 (d2) potential energy curves.

2. Orbital occupancy profiles for [2Fe-2S] cluster

In the main text we reported the average number of Fe electrons for the ground state of the [2Fe-2S] cluster and its two excited states S_1 and S_2 . Here, for completeness, we report in Fig. A4 the full occupancy profile **n** for each of these states. A representation of the orbitals can be found in [39, 42].

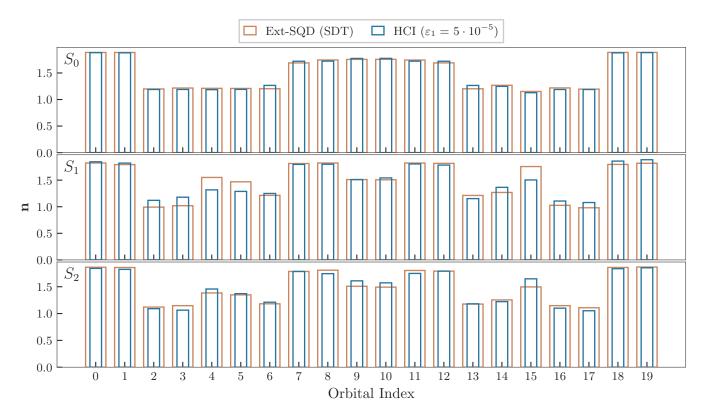


FIG. A4. [2Fe-2S] occupancy profiles. Orbital occupancy profiles \mathbf{n} for the ground state (S_0) and the first two singlet excited states $(S_1 \text{ and } S_2)$ of the [2Fe-2S] cluster. The occupancies obtained with Ext-SQD with the inclusions of singles, doubles, and triples are compared to the Heat-Bath CI results (see main text for more info on the methods).