Quantum Simulation of Ligand-like Molecules through Sample-based Quantum Diagonalization in Density Matrix Embedding Framework

Ashish Kumar Patra[1], **Anurag K. S. V.**[1], **Sai Shankar P.**[1], **Ruchika Bhat**[1, 2], **Raghavendra V.**[3], **Rahul Maitra**[4], **and Jaiganesh G.**[*, 1]

[1]Qclairvoyance Quantum Labs, Secunderabad, TG 500094, India.
[2]The University of Arizona, Tucson, AZ 85721, USA.
[3]SRM Institute of Science and Technology, Chennai, TN 603203, India.
[4]Indian Institute of Technology Bombay, Mumbai, MH 400076, India.

ABSTRACT

The accurate treatment of electron correlation in extended molecular systems remains computationally challenging using classical electronic structure methods. Hybrid quantum-classical algorithms offer a potential route to overcome these limitations; however, their practical deployment on existing quantum computers requires strategies that both reduce problem size and mitigate hardware noise. In this work, we investigate ground-state energy calculations of ligand-like molecules using Sample-based Quantum Diagonalization (SQD) within the Density Matrix Embedding Theory (DMET) framework, focusing on low-symmetry systems with diverse bonding motifs that exhibit subsystem-dependent variations in fragment-environment entanglement. These entanglement-based variations directly influence bath orbital construction, impurity sizes, and the structure of the embedded Hamiltonians, posing nontrivial challenges for both embedding and quantum sampling. By combining DMET fragmentation with SQD-based construction of reduced configuration spaces through quantum sampling and iterative configuration recovery, we perform quantum simulations on IBM's Eagle R3 (`ibm_sherbrooke`) and IBM's Heron R3 (`ibm_boston`) superconducting quantum hardware thereby, showing that the entanglement structure across embedding subsystems plays a central role in determining the efficiency and accuracy of the simulations. Despite these complexities, we show that the DMET-SQD framework yields ground-state energies in strong agreement with DMET-FCI benchmarks, achieving chemical accuracy (1 kcal/mol) across all systems studied. These results demonstrate that SQD-based quantum simulations can be robustly extended to low-symmetry, chemically realistic, industry relevant molecules, and highlight the importance of entanglement-aware embedding strategies for scalable quantum electronic structure calculations.

Keywords: Quantum Computing · Quantum Simulation · Quantum Chemistry · Sample-based Algorithms · Embedding Methods · Hybrid Quantum-Classical Algorithms

1 Introduction

In quantum chemistry, the accurate simulation of many-electron systems remains a central goal. However, the increasing complexity of molecular systems renders such endeavours computationally demanding, often necessitating varying levels of approximation[1]. Current classical computational resources are insufficient to fully capture electron correlation in larger molecular systems, such as proteins. More recently, distributed computing and an efficient flavor of Full Configuration Interaction (FCI) known as the Small-Tensor Product Distributed Active Space (STP-DAS) has been used to simulate the *HBrTe* molecule leading to the diagonalization of a Hilbert space exceeding 10^{15} determinants[2] involving 88 electrons and 100 spin orbitals in the x2c-TZVPall basis set[3]. For comparison, one of the largest conventional FCI simulations carried out using classical methods involved propane (C_3H_8), requiring the diagonalization of a Hilbert space spanning approximately 1.31×10^{12} determinants[4], even when employing a minimal STO-3G basis set[5]. In quantum chemistry, Hartree-Fock (HF) theory continues to serve as the foundational approximation for all subsequent methods[6].

With the advent of quantum computers, the accurate simulation of chemically and biologically relevant systems are expected to become increasingly tractable. Quantum algorithms promise the capability to model and evaluate protein-ligand interactions[7–9] and binding energetics, explore conformational landscapes of flexible bio-molecules[10–16]-tasks that rapidly become intractable on classical computers[17].This has motivated increasing interest in quantum computing, which offers alternative algorithmic frameworks capable of addressing many-electron correlations more efficiently[18,19].

*(Corresponding Author) email: jaiganesh@qclairvoyance.in, drjaiganesh15@gmail.com

Fault-tolerant Application-Scale Quantum (FASQ)[20] algorithms, such as Quantum Phase Estimation (QPE)[21] and Quantum Singular Value Transformation (QSVT)[22,23], face significant challenges due to their requirement for deep quantum circuits, which remain impractical on current Noisy Intermediate-Scale Quantum (NISQ) hardware[24]. To facilitate execution on near-term devices, variational approaches such as the Variational Quantum Eigensolver (VQE)[25] were proposed. VQE evaluates the expectation value of the Hamiltonian directly through a hybrid classical-quantum optimization loop, iteratively improving the overlap between the parametrized quantum state and the true ground state. Recent developments[26] have produced more hardware-efficient variants, yet key limitations persist, including statistical fluctuations and physical noise inherent to NISQ devices, as well as the requirement for prohibitively large numbers of samples and extensive error mitigation strategies[27].

Early hardware demonstrations of molecular problems using VQE, such as protein-ligand interactions, have been limited to active spaces involving around four qubits[7], primarily restricted to small organic systems. To overcome these limitations, sample-based approaches such as Sample-Based Quantum Diagonalization (SQD)[28] have emerged. SQD integrates Quantum Selected Configuration Interaction (QSCI)[29] with an error-mitigation scheme known as Iterative Self-Consistent Configuration Recovery (S-CoRe)[28] and employs classical Selected Configuration Interaction (SCI)[30] for post-processing diagonalization. These methods have enabled the quantum sampling of molecules with up to 36 spatial orbitals, including the [4Fe-4S] complex in a 77-qubit experiment, using 6400 nodes of the *Fugaku* super-computer[28]. We note that SQD and VQE are complementary rather than competing paradigms: whereas VQE performs iterative variational optimization of a parametrized ansatz, SQD leverages quantum sampling to construct a compact configuration subspace for subsequent classical diagonalization, with the dominant computational burden borne by the classical post-processing step.

While quantum computers are anticipated to handle strongly correlated systems more effectively, providing a compelling rationale for their adoption, counterarguments emphasize potential limitations[31]. Nonetheless, within the emerging paradigm of Quantum-Centric Super-Computing (QCSC)[28], where the samples obtained from the quantum computers are further processed in classical super-computers in a hybrid methodological framework, SQD has facilitated a series of further algorithmic advancements.

Recent developments in sample-based quantum algorithms include applications to materials simulation[32], excited-state energy estimation[33], hybrid methods combining SQD with Krylov Subspace Diagonalization[34] and QDrift[35], leading to techniques such as Sample-based Krylov Quantum Diagonalization (SKQD)[36], Quantum Krylov using Unitary Decomposition (QKUD)[37] and SqDRIFT[38]. Additionally, methods have been developed to iteratively expand the configuration space, achieving chemical accuracy while sampling only a fraction of the symmetry space (configurations obeying number and spin symmetries)[39–42], thereby reducing the classical computational costs as compared to SQD[28].

Classical fragmentation techniques offer a practical route to handling the exponential scaling of electronic structure methods by dividing large molecular systems into smaller subsystems. By decomposing the global wavefunction or Hamiltonian into fragment-level contributions, these approaches enable the study of chemically realistic systems without treating the entire system at once, while retaining a controlled description of inter-fragment interactions. Prominent examples include Density Matrix Embedding Theory (DMET)[43–46] and its multi-reference variants[47,48], Divide and Conquer (DC) method[49,50], and the Fragment Molecular Orbital (FMO) framework[51], all of which have demonstrated practical applicability across a wide range of molecular and materials problems[52]. These techniques provide a systematic route for reducing the effective dimensionality of quantum chemical calculations by localizing electronic structure in real space or orbital subspaces, while incorporating environmental effects through different mechanisms: embedding potentials in density-based approaches such as DFT embedding, density matrix matching in DMET, and effective Hamiltonians in wavefunction-based embedding frameworks[43,53].

Beyond these foundational approaches, several modern fragmentation schemes aim to preserve higher-order correlation effects while maintaining computational efficiency[54–60]. Coupled-cluster downfolding methods construct effective low-dimensional Hamiltonians by integrating out inactive degrees of freedom using classical CC theory, yielding size-consistent and systematically improvable reduced models[54,55,61]. Variational Localized Active Space methods, such as LASSCF[56,57], explicitly partition the active space into weakly entangled local subspaces that can be optimized independently, providing a natural interface for hybrid classical-quantum workflows. Related embedding paradigms include wavefunction-based embedding methods (EWF)[58], Many-Body Expansion (MBE)[59,60], and Dynamical Mean-Field Theory (DMFT)[62,63], which employ different strategies for capturing non-local correlation and dynamical effects. More recently, Bootstrap Embedding (BE)[64–67] has been introduced as a hybrid quantum-classical embedding framework that builds upon DMET by enforcing consistency of reduced density matrices across fragment boundaries, matching the density matrices of the central region of a fragment with those of its neighboring fragments.

In parallel, several approaches have focused on adapting classical fragmentation techniques for use with quantum simulation algorithms. Notable examples include DMET-VQE implementations on superconducting and trapped-ion platforms[7,15,68], as well as DMET-SQD[14] and EWF-SQD[69] approaches demonstrated on IBM quantum hardware. Related efforts such as divide-and-conquer VQE (DC-VQE)[70] and fragment molecular orbital extensions to quantum solvers, including Effective FMO-VQE[71], further illustrate the breadth of strategies combining classical fragmentation with quantum sampling or variational

techniques. Collectively, these works define a growing ecosystem of hybrid methods that aim to balance accuracy, scalability, and near-term hardware constraints in quantum-enabled electronic structure calculations.

SQD has gained significant interest in recent times[28,39–41] and is subject to further exploration to develop and assess its performance within fragmentation frameworks[14,69,72]. Towards this, we assess the feasibility of SQD within the DMET framework for ligand-like molecules, extending upon the recently introduced DMET-SQD framework[14] for strongly correlated symmetric systems. Herein, we extend this framework to accommodate a wider class of molecules that are biologically relevant, unlike the previously studied systems[14], which had high point group symmetry; these ligand-like molecules belong to low-symmetry classes (typically C_1). This makes the universal thresholding criteria to pick the dominant configurations over individual embedding problems a challenging task, a difficulty that provides a rigorous test of the robustness and practical applicability of the proposed approach.

The ligand-like molecules considered for the study include: Cyanic Acid($HOCN$), Formaldehyde Oxime (CH_3NO), O-methyl-hydroxylamine (CH_5NO), Methyl Isocyanate (C_2H_3NO), Acetaldehyde Oxime (C_2H_5NO), Carbamide/Urea (CH_4N_2O), Nitrosyl Chloride ($NOCl$), and Hydroxythiocyanate ($HOSCN$). All of these compounds have molecular weights below 76 Da (see Fig. 4) and together span a chemically diverse set of small, pharmacologically motivated molecular systems. Several members of this set, particularly urea derivatives and methoxyamine-related systems, have established relevance in medicinal chemistry[73–76]. Furthermore, due to the heterogeneous nature of individual embedding problems, unlike previous studies, the variation in the fragment-environment entanglement across the studied low-symmetry molecules manifests directly in the bath orbital construction and impurity sizes, posing nontrivial challenges for both embedding and quantum sampling. This will be further elaborated in the Results Section (Section 4), and makes the chosen set of molecules an appropriate testbed for assessing DMET-SQD in the context of biologically and industrially relevant molecular systems. In particular, the virtual orbital occupation threshold ε_{occ} governing bath orbital inclusion emerges as a critical, system-dependent parameter whose selection directly determines the balance between embedding accuracy and quantum hardware feasibility, and we therefore subject it to a systematic sensitivity analysis.

Concisely detailing the work, in **Section 2** an overview of the methods used are described briefly, in **Section 3** the experimental details and methodology are discussed where we emphasize that the variable entanglement structure of the embedded fragments may vary the performance of SQD due to the lack of symmetry in the molecules, in **Section 4** the results obtained are discussed and finally in **Section 5** we conclude the work and mention future prospects.

2 Theoretical Background

2.1 Hamiltonian Formulation and Local Unitary Cluster Jastrow ansatz

Under the Born-Oppenheimer approximation[77], the electronic structure problem is most conveniently expressed in second quantization, where the Hamiltonian is written in terms of fermionic creation and annihilation operators acting on a basis of spin orbitals[78]:

$$\hat{H}_{sq} = \sum_{i,j} h_{ij} \hat{a}_i^\dagger \hat{a}_j + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s. \tag{1}$$

This representation separates one-electron and two-electron contributions and forms the natural interface for quantum simulation, since it can be systematically mapped onto qubit operators using standard fermion-to-qubit encodings such as the Jordan-Wigner transformation[79]. In this work, we initialize the quantum register in the restricted Hartree-Fock (HF) reference state, which is a single Slater determinant that is straightforward to prepare on quantum hardware and typically exhibits a significant overlap with the true ground state.

Starting from this reference, we construct a correlated variational wavefunction using the Local Unitary Cluster Jastrow (LUCJ) ansatz[80], which is designed to balance chemical expressivity with circuit efficiency on near-term quantum devices. The ansatz parameters are initialized using amplitudes inspired by coupled-cluster singles and doubles (CCSD) theory[81,82], providing a chemically motivated starting point for optimization. The LUCJ ansatz consists of multiple layers, each comprising an orbital-rotation operator, a diagonal Jastrow interaction that captures electron-electron correlations, and the inverse orbital rotation, thereby incorporating correlation effects in a structured and hardware-aware manner[83,84]. The resulting variational state is given by

$$|\psi\rangle_{UCJ} = \prod_{\mu=1}^{L} e^{\hat{K}_\mu} e^{i\hat{J}_\mu} e^{-\hat{K}_\mu} |\psi\rangle_{HF}\,. \tag{2}$$

Here, $|\psi\rangle_{HF}$ denotes the Hartree-Fock reference state and L is the total number of LUCJ layers. The operators $\hat{K}_\mu$ and $\hat{J}_\mu$ respectively generate orbital rotations and two-body correlation effects within each layer, enabling the ansatz to systematically incorporate electronic correlation beyond the mean-field level while maintaining a circuit structure amenable to near-term quantum hardware. For closed-shell systems, a spin-balanced form of the ansatz is employed to enforce spin symmetry and reduce the number of independent variational parameters, improving optimization stability without sacrificing expressivity.

2.2 Density Matrix Embedding Theory

Density Matrix Embedding Theory (DMET)[44] is a frequency-independent quantum embedding method conceptually related to Dynamical Mean-Field Theory (DMFT)[62], in which the environment of a chosen fragment is represented through a set of entangled orbitals, referred to as bath orbitals. In DMET, a set of user-defined fragment orbitals A_y is selected from a localized orbital basis, and the remaining orbital space constitutes the environment Env_y. This environment is decomposed into entangled bath orbitals and unentangled core and virtual orbitals according to

$$Env_y = B_y \cup Cor_y \cup Vir_y, \tag{3}$$

where B_y denotes the bath orbitals capturing fragment-environment entanglement, while Cor_y and Vir_y correspond to unentangled occupied and virtual orbitals, respectively.

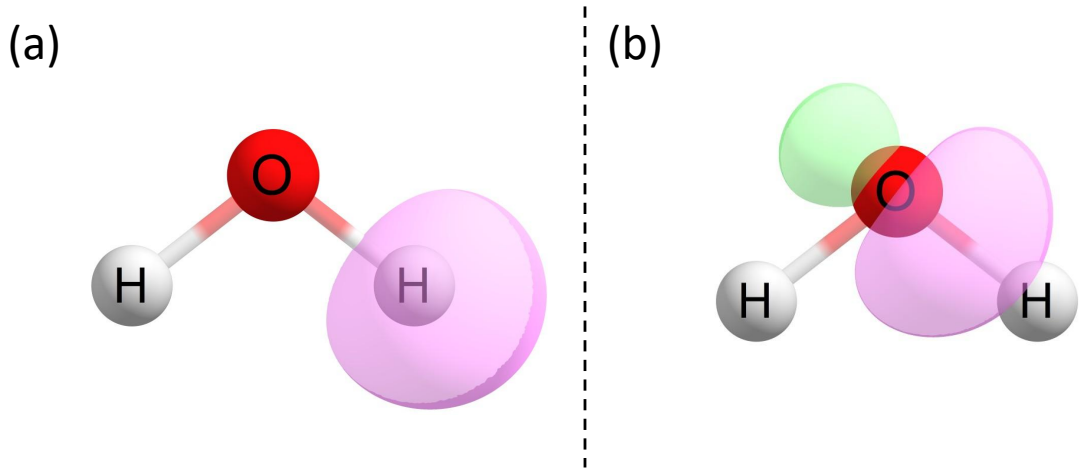

Figure 1. Illustrative example of bath-orbital construction for the H_2O molecule in the STO-3G basis. The system is fragmented so that the orbital localized on a single hydrogen atom constitutes the fragment. **(a)** shows an iso-surface of this localized fragment orbital; **(b)** shows the corresponding bath orbital obtained via DMET, which is a linear combination of the remaining localized spatial orbitals in the molecular orbital basis. Together, **(a)** and **(b)** span the impurity subspace for the H-fragment.

The bath orbitals are constructed from a mean-field reference (typically Hartree-Fock) by projecting the one-particle reduced density matrix (1-RDM) onto the fragment subspace and diagonalizing the resulting matrix. The eigenvectors with fractional occupations define the bath orbitals, and their number is bounded by the number of fragment orbitals[85,86]. An illustrative example of bath construction for a hydrogen fragment in a water molecule is shown in Fig. 1. The fragment and bath orbitals together form the impurity space on which an embedded Hamiltonian is solved to obtain the correlated wavefunction. A global chemical potential μ_{glob} is adjusted to ensure electron number consistency across all fragments, corresponding to a one-shot interacting bath formulation[43].

The fragment energy E_{A_y} is given by

$$E_{A_y} \approx \sum_{p\in A_y} \left(\sum_{q}^{L_{A_y}+L_{B_y}} \left(\frac{t_{pq}+\bar{h}^y_{pq}}{2} \right) D^y_{qp} + \frac{1}{2} \sum_{qrs}^{L_{A_y}+L_{B_y}} (pq|rs) P^y_{qp|sr} \right), \tag{4}$$

where y indexes the fragments in the partitioning of the molecular system, and E_{A_y} denotes the energy contribution associated with fragment A_y. The indices p,q,r,s label spatial orbitals belonging to the impurity space, defined as the union of fragment and bath orbitals with total dimension $L_{A_y}+L_{B_y}$. The matrix elements t_{pq} denote the one-electron integrals in the molecular orbital basis, while $\bar{h}^y_{pq}$ are effective one-body terms incorporating the mean-field contribution of the core orbitals. The

quantities $D^y_{qp} = \langle \hat{a}^\dagger_p \hat{a}_q \rangle$ and $P^y_{qp|sr} = \langle \hat{a}^\dagger_p \hat{a}^\dagger_r \hat{a}_s \hat{a}_q \rangle$ denote the fragment one- and two-particle reduced density matrices (1-RDM and 2-RDM), respectively. The two-electron integrals $(pq|rs)$ are Coulomb integrals in the molecular orbital basis. The prefactors of $1/2$ avoid double counting when summing over all orbital indices, and the total energy of the system is obtained by summing the contributions from all fragments.

2.3 Sample-based Quantum Diagonalization

SQD[28] is a combination of the quantum sampling method QSCI[29] and an error-mitigation technique introduced in the work[28], S-CoRe, which "recovers" the noisy configurations. These configurations thus obtained create a sub-space, which is further expanded, and then the projection sub-space is obtained over which SCI[30] is performed to obtain the ground state energy.

2.3.1 Quantum State Preparation and Sampling

We begin with the goal of approximating the ground state energy of a many-body Hamiltonian by constructing and (classically) diagonalizing a reduced Hilbert space that contains the *dominant configurations* contributing to the ground state configuration, *sampled using a quantum computer*. Starting from an approximate ground state, like $|\psi\rangle_{HF}$, a variational ansatz is initialized with suitable parameters: $|\psi\rangle_{init} = \hat{U}(\theta_1, \theta_2, ..., \theta_n)\,|\psi\rangle_{HF}$ (in our case, the LUCJ ansatz with CCSD parameters) to increase the overlap between the trial state with the true ground state, as provided in Fig. 2.

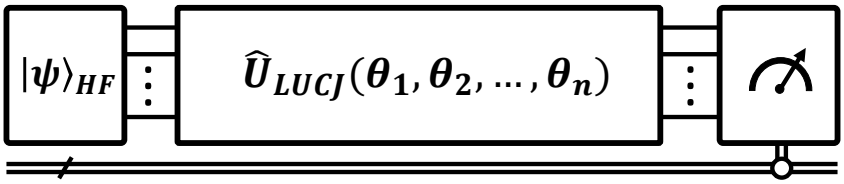

Figure 2. Quantum Circuit Diagram for performing Sample-based Quantum Diagonalization. The initial state is prepared as the Hartree-Fock configuration, and is acted upon by an LUCJ ansatz which is pre-initialized with t^a_i and t^{ab}_{ij} parameters obtained through CCSD calculation. Finally, measurement is performed in the $\hat{\sigma}_z$ basis.

Quantum sampling is performed in the computational basis, and a set $\mathscr{S}_{samp}$ is obtained with d configurations:

$$\mathscr{S}_{samp} = \{|\phi_1\rangle, |\phi_2\rangle, ..., |\phi_d\rangle\} \tag{5}$$

where every $|\phi_i\rangle$ can be represented as a tensor product of the α spin and β spin configurations: $|\phi_i\rangle = |\alpha_i\rangle \otimes |\beta_i\rangle$. The total number of electrons in the molecular systems is designated as N_{elec}, and correspondingly the number of electrons in α configuration and β configurations are designated as N_α and N_β respectively. For molecular systems which contain fully-filled spatial orbitals (closed-shell) and which possess zero spin, the following conditions must be satisfied for all the computational basis states contributing to the eigen-states of the system:

$$N_\alpha + N_\beta = N_{elec}; \quad N_\alpha - N_\beta = 0. \tag{6}$$

Due to the noisy nature of quantum computers, we end up also sampling configurations which violate the particle-number and spin-z symmetries mentioned in Eq.(6). These "noisy" configurations are recovered using configuration recovery[28].

2.3.2 Self-Consistent Iterative Configuration Recovery

S-CoRe, or Configuration Recovery, is an iterative procedure designed to restore the sampled noisy configurations. Given an approximate ground-state occupation number distribution for the state-vector $|\Psi\rangle$:

$$n_{p\sigma} = \langle \Psi | \hat{a}^\dagger_{p\sigma} \hat{a}_{p\sigma} | \Psi \rangle, \tag{7}$$

where $n_{p\sigma}$ is the occupancy corresponding to the p^{th} spatial orbital with σ spin configuration, S-CoRe updates each sampled configuration, over multiple batches (the b-th batch indexed by $b \in \{1, \ldots, K\}$, where K is the total number of batches) by probabilistically flipping occupations of the spin-orbitals based on the values of $n_{p\sigma}$ until the total electron number and spin projection match the target values. This produces a set of corrected configuration sub-spaces $\mathscr{S}^{(b)}_{CR}, \forall b \in \{1, \ldots, K\}$ that respect the intended symmetries while remaining close to the original sampled distribution.

For projection and diagonalization, the sub-spaces $\mathscr{S}^{(b)}_{CR}$ are converted to $\mathscr{S}^{(b)}_{proj}$ as detailed in Section 2.3.3. With the $E^{(b)}$ and $|\Psi^{(b)}\rangle$ obtained post the classical diagonalization for all the K batches, the lowest energy $min_b E^{(b)}$ is taken as the current

best estimate, and the average orbital occupation is thus calculated as:

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

which is again used in the next iteration and so on. This self-consistent iteration ensures that the symmetry constraints are enforced, and the recovered sub-space contains configurations closer to the desired orbital occupation distribution. From these recovered configurations at each iteration, K batches of subspaces $\{\mathscr{S}^{(b)}_{CR}\}_{b\in K}$ are obtained.

2.3.3 Selected Configuration Interaction

For the set of configurations obtained in the b^{th} batch, the sub-space is designated as $\mathscr{S}^{(b)}_{CR}$. As the subsequent steps remain the same for all the batches, we shall drop the super-script and denote the sub-space as $\mathscr{S}_{CR}$ without loss of generality.

In SQD, the configurations $\mathscr{S}_{CR} = \{|\alpha_1\rangle \otimes |\beta_1\rangle, |\alpha_2\rangle \otimes |\beta_2\rangle, ..., |\alpha_d\rangle \otimes |\beta_d\rangle\}$ do not represent the sub-space onto which the Hamiltonian is projected. Instead in the case of **closed-shell systems** all the unique alpha configurations $\mathscr{A}_{unique} = \{|\alpha_1\rangle, |\alpha_2\rangle, ..., |\alpha_d\rangle\}$ and unique beta configurations $\mathscr{B}_{unique} = \{|\beta_1\rangle, |\beta_2\rangle, ..., |\beta_d\rangle\}$ are again made into a final set:

$$\mathscr{P}_{unique} = \mathscr{A}_{unique} \cup \mathscr{B}_{unique} \tag{9}$$

Now, a new configuration space is obtained from the tensor product of $\mathscr{P}_{unique}$ with itself, which we define as $\mathscr{S}_{proj}$:

$$\mathscr{P}_{unique} \otimes \mathscr{P}_{unique} = \mathscr{S}_{proj} = \{\, |\lambda\rangle \otimes |\phi\rangle : \forall\ |\lambda\rangle \in \mathscr{P}_{unique},\ \forall\ |\phi\rangle \in \mathscr{P}_{unique} \,\}. \tag{10}$$

Hence, $\mathscr{S}_{proj}$ is the **sub-space onto which the Hamiltonian is projected onto, and not the** $\mathscr{S}_{CR}$ obtained after configuration recovery. $\mathscr{S}_{proj}$ is subsequently diagonalized using Davidson's Algorithm[87]. The construction in which the unique configurations from $\mathscr{A}_{unique}$ and $\mathscr{B}_{unique}$ are merged into $\mathscr{P}_{unique}$, followed by forming the tensor-product space $\mathscr{P}_{unique} \otimes \mathscr{P}_{unique}$, is motivated by the observation that the ground-state wavefunction generally exhibits correlated contributions from multiple combinations of spin-resolved configurations.In particular, if a configuration of the form $|\alpha\rangle \otimes |\beta\rangle$ carries a non-zero amplitude in the ground state, then configurations formed from the same α and β sectors in different pairings may also contribute, as observed in correlated wavefunction structures[39]. This construction is also consistent with the original SQD formulation, where the subspace is generated by combining unique spin configurations to form all possible pairs[28].

The Hamiltonian is projected (for every b^{th} sub-space) into the sub-space $\mathscr{S}_{proj}$,

$$\hat{H}_{\mathscr{S}_{proj}} = \hat{P}_{\mathscr{S}_{proj}} \hat{H} \hat{P}_{\mathscr{S}_{proj}}, \qquad \hat{P}_{\mathscr{S}_{proj}} = \sum_{\mathbf{x}\in\mathscr{S}_{proj}} |\mathbf{x}\rangle\langle\mathbf{x}|, \tag{11}$$

and the solution is obtained by performing diagonalization using Davidson's method[87], finally providing us $E^{(b)}$ and $|\Psi^{(b)}\rangle$, which are again utilized in Eq. (8) to obtain the average orbital occupancies.

Because the diagonalization cost scales polynomially in $|\mathscr{S}_{proj}|$, the method remains tractable as long as the number of selected configurations does not grow exponentially with system size[87].

3 Methodology

This section describes the DMET-SQD framework, integrating fragmentation, impurity solving, and self-consistent quantum procedures via global chemical potential. Section 3.1 details the workflow; Section 3.2 covers hardware, software, and settings.

3.1 DMET-SQD Workflow

The workflow[15] used to perform DMET-SQD is as follows:

1. **Full System Low-level Reference:** The molecular co-efficient matrix C, and the HF determinant (in the case of single-reference systems) of the whole system are obtained using Restricted HF method (here we only encounter molecular systems with fully filled spatial orbitals).
2. **Orbital Localization and Fragmentation:** Then, the molecular spatial orbitals are localized (using Meta-Löwdin localization[88]) and based on the user-defined criteria (in this case, one atom per fragment), the orbitals are fragmented.

3. **Bath Orbital Construction:** The bath orbitals are constructed with the remaining environment orbitals for the fragment A_y. Here, the occupation threshold ε_{occ} is set to 10^{-13} for discarding virtual orbitals [43].
4. **Impurity Fock Matrix Updation:** The Fock matrix $\hat{F}_{imp}$ (where $imp = A_y + B_y$) for each of the impurities is obtained, and the chemical potential μ_{glob} is subtracted from the diagonal entries corresponding to the fragment orbitals:

$$\hat{F}^{new}_{imp}(\mu_{glob}) = \hat{F}_{imp} - \mu_{glob}\hat{P}_{frag}, \quad \hat{P}_{frag} = \begin{cases} 1 & \text{if } i = j \text{ and } i \in A_y \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

5. **Impurity SCF Updation:** Using $\hat{F}^{\text{new}}_{imp}(\mu_{\text{glob}})$ from Eq. (12), an SCF calculation is carried out in the impurity space to obtain updated molecular properties, such as the impurity coefficient matrix $C^{\text{new}}_{imp}(\mu_{\text{glob}})$ and the corresponding one-particle density matrix $D^{y}(\mu_{\text{glob}})$. The standard convention is to initialize $\mu^{init}_{glob} = 0$.
6. **High-level Quantum Solver (SQD):** Now, we use SQD, *as the high-level solver* to estimate the ground state energy and eigen-vector of these "impurities". The one-electron ($h^{pq}_{imp}(\mu_{glob})$) and two-electron ($h^{pqrs}_{imp}(\mu_{glob})$) Hamiltonian terms (where $pqrs \in imp$) obtained after the localization, serve as the $\hat{H}_{emb}(\mu_{glob})$, and the configurations obtained through quantum sampling (and subsequent Configuration Recovery) are used to form the subspace onto which $\hat{H}_{emb}$ is projected. *In the case of calculating the reference energies, this step would be replaced by FCI, without any loss of generalization.*
7. **Fragment Energy and RDM Extraction:** Using the eigen-vector obtained in the previous step, the following quantities are calculated:
   - fragment energy $E_{A_y}(\mu_{glob})$ using Eq. (4),
   - number of electrons in the fragment[15] $N^{ele}_{A_y}$, where $N^{ele}_{A_y}(\mu_{glob}) = \sum_{p \in A_y} D^{y}_{pp}(\mu_{glob})$
   - 1-RDM $D^{y}_{pq}(\mu_{glob})$ and
   - 2-RDM $P^{y}_{qp|sr}(\mu_{glob})$ of the impurities, where $pqrs \in imp$.
8. **Aggregation of Fragment Contributions:** Now similar to the fragment A_y, the steps 3 to 7 are repeated for all the remaining $N_{frag} - 1$ fragments, where N_{frag} represents the number of fragments in the system. The total energy $E_{tot}(\mu_{glob})$, the number of electrons of the system N_{tot}, and the error in electron number N^{ele}_{err} can be thus obtained by:

$$E_{tot}(\mu_{glob}) = E_{nuc} + \sum_{y}^{N_{frag}} E_{A_y}(\mu_{glob}), \quad N_{tot}(\mu_{glob}) = \sum_{y}^{N_{frag}} N^{ele}_{A_y}(\mu_{glob}), \quad N^{ele}_{err}(\mu_{glob}) = N_{tot}(\mu_{glob}) - N_{true} \tag{13}$$

   where E_{nuc} is the nuclear repulsion energy of the total system, and N_{true} is the exact number of electrons in the system (which we know a priori).
9. **Chemical Potential Optimization:** Next, the root of the function $N^{ele}_{err}(\mu_{glob})$ is computed to within the convergence threshold ε_{conv}, by iteratively updating μ_{glob}, and repeating the steps 3 to 8, using a root finding method (in this case we use Newton-Secant method[89]). In other words, we wish to obtain a $\mu^{\star}_{glob}$ such that:

$$\left|N_{tot}(\mu^{\star}_{glob}) - N_{true}\right| < \varepsilon_{conv}. \tag{14}$$

The nine workflow steps described above are grouped into six broader conceptual categories, designated as **Parts I-VI**, to provide a higher-level organizational structure for the DMET-SQD algorithm, as illustrated in Fig. 3. **Step 1**, **Step 2** can be considered to fall under a common section, designated as **Part I**, which deals with initial low-level solving of the entire system to obtain certain pre-requisites, **Steps 3, 4** and **5**, deal with the creation of bath orbitals, and the inclusion of μ_{glob} into the hamiltonians of the impurities, which can be designated as **Part II**. **Step 6** falls under **Part III**, where the previously obtained Hamiltonians are solved using a high-level solver, such as SQD. **Step 7** involves obtaining the relevant information from the state-vector obtained from high-level solving. This crucial step allows us to perform further iterations, and can be designated as **Part IV**. **Step 8** and **Step 9** involve looping over the fragments to obtain the total energy and the number of electrons in the system, and subsequently tweaking μ_{glob} to satisfy Eq. (14), respectively. These steps are designated as **Part V** and **Part VI**.

3.2 Experimental Details

The sampling part of the quantum simulation experiments in this work was executed on the Eagle R3 quantum processor (IBM Sherbrooke, containing 127 qubits)[84]. The corresponding hardware calibration metrics recorded immediately prior to the computation are provided in Section C of the Supplementary Material. These values represent the initial calibration state of the device and should not be interpreted as a faithful reflection of the noise characteristics during execution. This distinction is

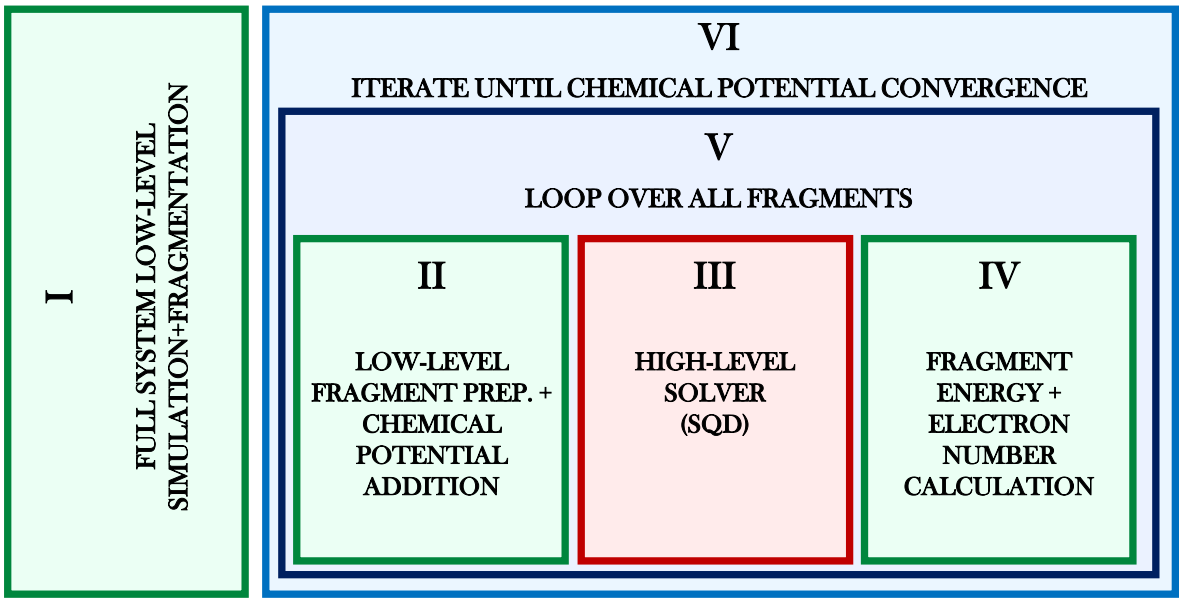

Figure 3. A simplified workflow for DMET-SQD

important, as the full set of molecular simulations required approximately two days to complete, largely due to queueing delays between successive quantum-sampling jobs, during which hardware performance may drift inevitably due to queue wait times on shared quantum hardware, with drift bounded by the ∼24-hour automated recalibration cycle as noted in Section C of the Supplementary Material.

For DMET-SQD and DMET-FCI, `Tangelo v0.4.3`[90] was used. LUCJ circuits were generated using `ffsim v0.10.0`[91], and quantum circuits were constructed and transpiled with `Qiskit v1.4.2`[92] and `Qiskit IBM Runtime v0.36.1`. All quantum-chemistry tasks (HF, CCSD, and SCI) were carried out using `PySCF v2.10.0`[88]. The root-finding method used was the Newton–Secant method from `scipy v1.15.3`[93].

The molecules considered in this study are listed in Section B of the Supplementary Material. The STO–3G basis set was employed to enable controlled benchmarking of the DMET-SQD framework against DMET-FCI, allowing for direct and systematic comparison in a regime where exact classical reference solutions remain tractable. In particular, the use of a minimal basis allows us to avoid additional approximations such as active space truncation in the DMET-FCI calculations, thereby ensuring a consistent and unbiased reference. Furthermore, the reduced Hilbert space complexity makes this setting well suited for studying the behavior of the embedding and sampling procedures under NISQ-era hardware constraints. Fragmentation was performed such that each atom is treated as an individual fragment for all molecules.

While ROHF-based DMET formulations for open-shell systems are well-established[94–96], the ligand-like molecules considered in this study are closed-shell systems, and a restricted Hartree-Fock reference is therefore employed for constructing the DMET bath. The fragmentation scheme adopted is one-atom-per-fragment as a stringent test of bath orbital construction, including chemically nontrivial cases where covalent bonds, such as double bonds, are cut across fragments. This aggressive partitioning may introduce significant errors, but those errors will be identical in both DMET-SQD and DMET-FCI and can hence be safely ignored, while simultaneously keeping the sample space as small as possible to reduce shot budget for SQD sampling and also to reduce the classical post-processing overhead. We note that DMET is not rigorously size-consistent in the formal sense; however, since both DMET-SQD and DMET-FCI are evaluated under identical embedding assumptions, this limitation does not affect the conclusions drawn from their comparison.

Since the SQD solver relies on SCI for diagonalization, it is formally neither size-consistent nor size-extensive, and its accuracy depends explicitly on the configurations retained. However, the hardware-driven sampling and subsequent configuration recovery ensure that the most dominant determinants across the full Hilbert space are naturally picked, and this dominant subspace can be systematically enlarged to improve the accuracy. This ensures that despite the formal lack of size-consistency, the numerical accuracy is maintained even for larger basis sets and more extended systems within the scope of the approach considered here.

For benchmarking, FCI is employed as it accesses the full configuration space and is both size-consistent and size-extensive. This ensures that any deviation in correlation energy arises solely from the embedding and sampling procedure, thereby providing a rigorous validation of the ability of DMET-SQD to reproduce DMET-FCI correlation within the scope of the present study.

The number of shots used for quantum sampling was 10^4 and was kept consistent across all the impurities, regardless of the number of spin orbitals within them. For each impurity, 5 iterations of S-CoRe were performed. The maximum allowed samples per batch was set to 10^5, which is the quantum sub-sampling threshold that is applied to the configurations obtained post configuration recovery. The circuits were transpiled[97] with optimization set to level 3, and no additional error handling techniques were employed in these series of experiments. The convergence threshold ε_{conv} for μ_{glob} was set to 1.48×10^{-8}, which is the default parameter in `scipy.optimize.newton` method. The default virtual orbital threshold ε_{occ} was 10^{-13} for bath orbital construction. This threshold determines the inclusion of the non-negligibly entangled environment orbitals into the impurity space, thereby directly controlling the impurity size. Consequently, it influences both the amount of correlation captured and the quantum resource requirements of the corresponding SQD simulations.

3.3 Bath-Orbital Threshold Sensitivity Analysis

The construction of the DMET bath orbitals depends explicitly on the occupation threshold ε_{occ}, which determines whether a given environment orbital is classified as entangled and therefore included in the impurity space. Since the size of the resulting impurity Hamiltonian directly affects both the classical and quantum computational costs, it is important to assess the sensitivity of the embedding procedure to the choice of ε_{occ}.

To investigate this dependence, additional calculations were performed using a range of occupation thresholds,

$$\varepsilon_{\text{occ}} \in \left\{10^{-5}, 10^{-7}, 10^{-9}, 10^{-11}, 10^{-13}, 10^{-15}\right\}, \tag{15}$$

for representative embedding problem, particularly for the *HOSCN* molecule, in the Heron R3 quantum processor (IBM Boston), the hardware calibration details of which are provided in Section D of the Supplementary Material. For each threshold, the corresponding bath-orbital dimensions, impurity sizes, quantum resources, and the resulting energetics were analyzed. Particular attention was paid to the variation in fragment-environment entanglement across the studied low-symmetry molecules, since weakly entangled orbitals near the threshold can lead to significant changes in the impurity dimension.

The threshold sensitivity analysis was primarily performed to quantify the relationship between fragment-environment entanglement and impurity construction, and to determine whether the inclusion of weakly entangled bath orbitals substantially affects the resulting embedded Hamiltonians. Unless otherwise stated, all production DMET-SQD and DMET-FCI calculations reported in this work employ the value $\varepsilon_{\text{occ}} = 10^{-13}$, which was chosen to ensure that all non-negligibly entangled environment orbitals are retained in the impurity space while maintaining a tractable computational cost.

4 Results

This section is organized as follows: Section 4.1 presents the main accuracy and convergence results across all eight molecules; Section 4.2 examines the effect of the occupation threshold ε_{occ} on embedding and quantum resources for HOSCN; and broader implications of the results.

4.1 DMET-SQD Accuracy and Convergence

The absolute energy differences between the DMET-SQD energies and DMET-FCI energies of the molecules are presented in Fig. 4. As is evident, the energy differences obtained are consistently lower than 10^{-5} Ha, and in some cases are within micro-Hartree precision in comparison to DMET-FCI. We note that since DMET-SQD and DMET-FCI employ identical fragmentation settings, basis sets, and occupation thresholds, all systematic errors originating from basis set incompleteness, one-atom-per-fragment partitioning, and bath truncation cancel exactly in ΔE, which therefore isolates the accuracy of the SQD impurity solver relative to the FCI reference within an identical embedding framework, making this a stringent comparison of the chosen high-level quantum computing solver against the exact classical reference. Two additional factors that may influence the reported accuracy are the SQD subspace thresholding and hardware noise. The maximum number of samples per batch was set to 10^5 across all fragments, ensuring that SQD's quantum sub-sampling threshold does not constitute a limiting factor in the reported accuracy. The hardware calibration metrics recorded immediately prior to execution are provided in Appendix C.

The details pertaining to the fragmentation settings, quantum resource requirements, and the dimensions of the symmetry space and Hilbert space are summarized in Table 1. The QPU time 'T(s)' in the table is measured via the `IBM Qiskit Runtime` API's provided metadata and represents the cumulative quantum sampling time across all fragments and DMET iterations, excluding classical post-processing (S-CoRe and SCI diagonalization). The resource estimation is presented in the IBM Sherbrooke superconducting hardware basis gate set $\{X, R_z, \sqrt{X}(S_x), ECR\}$. The largest quantum-sampling experiment

Table 1. Quantum resource estimation summary for the molecular fragments. Beginning from the left: Chemical formula of the molecule, Fragment, number of spatial orbitals and number of electrons represented as (o, e), number of qubits used, depth of the quantum circuit, number of R_z, $\sqrt{X}(S_x)$, X, ECR gates used respectively for the specific instance of execution, dimension of the Symmetry Space(|S.S|), dimension of the Hilbert Space(|H.S|), the QPU time taken for performing DMET-SQD, and the number of iterations for μ_{glob} convergence (N_{it}).

Molecule	Frag	(o, e)	Qubits	Depth	R_z	S_x	X	ECR	|S.S|	|H.S|	T(s)	N_{it}
$HCNO$	[H]	(2, 2)	4	34	46	38	0	12	4	16	68	4
	[C]	(10, 10)	20	389	1939	1524	108	584	63,504	1,048,576		
	[N]	(10, 10)	20	365	1833	1548	88	550	63,504	1,048,576		
	[O]	(10, 10)	20	343	1804	1410	126	550	63,504	1,048,576		
CH_3NO	[C]	(10, 10)	20	366	1847	1454	134	580	63,504	1,048,576	107	4
	[H]	(2, 2)	4	34	43	38	0	12	4	16		
	[H]	(2, 2)	4	34	44	38	0	12	4	16		
	[H]	(2, 2)	4	35	47	38	0	12	4	16		
	[N]	(10, 10)	20	341	1814	1511	88	542	63,504	1,048,576		
	[O]	(10, 10)	20	378	1893	1479	138	586	63,504	1,048,576		
CH_5NO	[C]	(10, 10)	20	369	1946	1578	109	586	63,504	1,048,576	130	4
	[H]	(2, 2)	4	35	45	38	0	12	4	16		
	[H]	(2, 2)	4	35	48	38	0	12	4	16		
	[H]	(2, 2)	4	34	44	38	0	12	4	16		
	[H]	(2, 2)	4	35	46	38	0	12	4	16		
	[H]	(2, 2)	4	34	46	38	0	12	4	16		
	[O]	(10, 10)	20	369	1796	1463	120	552	63,504	1,048,576		
	[N]	(10, 10)	20	410	1950	1543	130	590	63,504	1,048,576		
C_2H_3NO	[C]	(10, 10)	20	591	1929	1597	126	604	63,504	1,048,576	130	4
	[C]	(10, 10)	20	396	1871	1503	125	578	63,504	1,048,576		
	[H]	(2, 2)	4	35	46	38	0	12	4	16		
	[H]	(2, 2)	4	35	46	38	0	12	4	16		
	[H]	(2, 2)	4	34	43	38	0	12	4	16		
	[N]	(10, 10)	20	371	1901	1532	104	568	63,504	1,048,576		
	[O]	(10, 10)	20	383	1868	1544	116	572	63,504	1,048,576		
C_2H_5NO	[C]	(10, 10)	20	361	1927	1589	107	580	63,504	1,048,576	165	4
	[C]	(10, 10)	20	343	1813	1432	134	568	63,504	1,048,576		
	[H]	(2, 2)	4	35	47	38	0	12	4	16		
	[H]	(2, 2)	4	35	47	38	0	12	4	16		
	[H]	(2, 2)	4	35	46	38	0	12	4	16		
	[H]	(2, 2)	4	35	48	38	0	12	4	16		
	[H]	(2, 2)	4	35	48	38	0	12	4	16		
	[N]	(10, 10)	20	375	1906	1545	94	568	63,504	1,048,576		
	[O]	(10, 10)	20	635	1865	1424	139	572	63,504	1,048,576		
CH_4N_2O	[C]	(10, 10)	20	633	1809	1405	142	568	63,504	1,048,576	119	4
	[N]	(10, 10)	20	655	1877	1389	131	588	63,504	1,048,576		
	[H]	(2, 2)	4	34	43	38	0	12	4	16		
	[H]	(2, 2)	4	35	47	38	0	12	4	16		
	[H]	(2, 2)	4	34	44	38	0	12	4	16		
	[H]	(2, 2)	4	33	42	38	0	12	4	16		
	[N]	(10, 10)	20	408	1925	1500	117	578	63,504	1,048,576		
	[O]	(10, 10)	20	444	1811	1437	122	564	63,504	1,048,576		
$NOCl$	[N]	(8, 10)	16	223	897	628	88	300	3,136	65,536	62	4
	[O]	(8, 10)	16	219	891	636	80	300	3,136	65,536		
	[Cl]	(12, 18)	24	342	2441	2117	109	738	48,400	16,777,216		
$HOSCN$	[H]	(2, 2)	4	34	43	38	0	12	4	16	98	4
	[O]	(10, 10)	20	454	1838	1427	199	619	63,504	1,048,576		
	[S]	(15, 18)	**30**	**1,081**	4384	4018	205	1351	**25,050,025**	**1,073,741,824**		
	[C]	(10, 10)	20	651	1911	1534	146	600	63,504	1,048,576		
	[N]	(10, 10)	20	558	2021	1622	143	646	63,504	1,048,576		

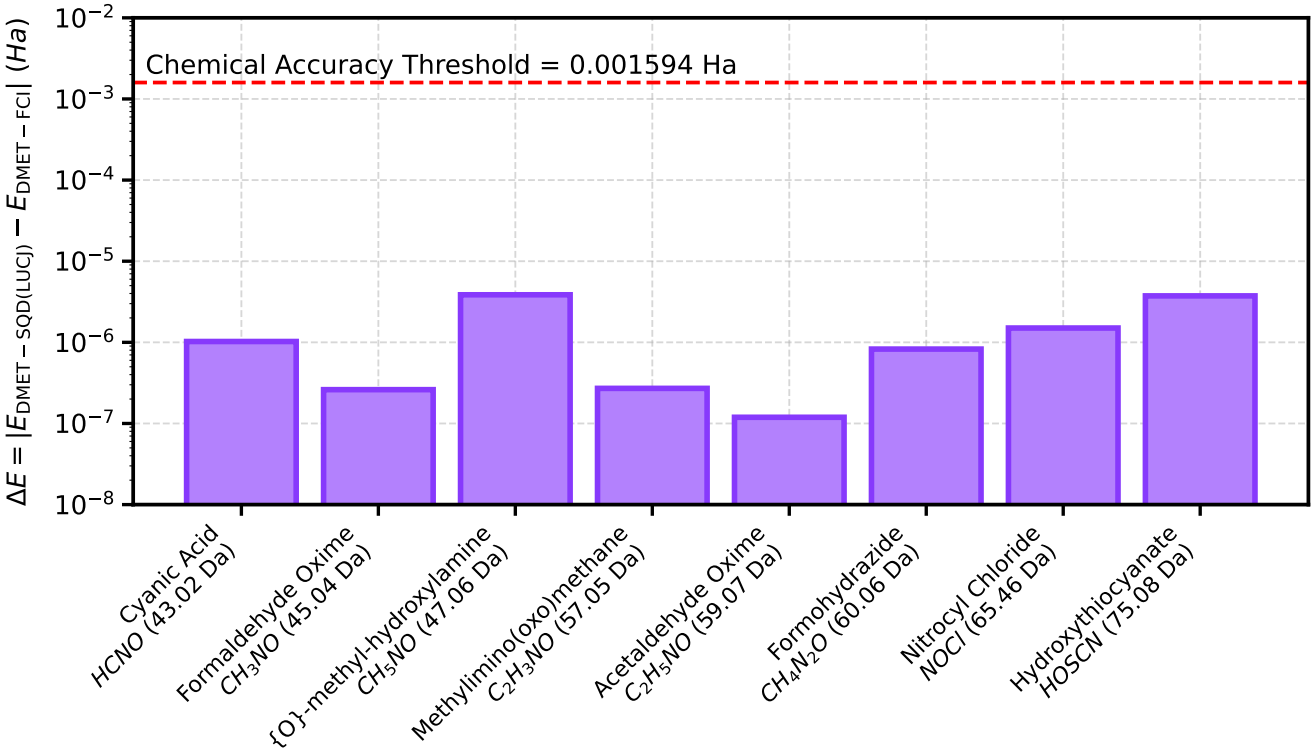

Figure 4. Comparison of the DMET-SQD energies obtained for the set of studied molecules on IBM Sherbrooke Quantum Hardware ($4^{th} - 5^{th}$ June, 2025) with DMET-FCI. All the energies obtained are within the chemical accuracy criterion (1 kcal/mol $\approx$ 0.001594 Ha), which is marked with the red-dashed line. The x-axis denotes the name of the studied molecule, the molecular formula, and the molecular weight. The y-axis denotes the absolute energy difference between the energy obtained through quantum hardware using DMET-SQD ($E_{DMET-SQD}$) and the reference energy $E_{DMET-FCI}$, which is defined as $\triangle E$. Since DMET-SQD and DMET-FCI employ identical fragmentation settings, basis sets, and occupation thresholds, ΔE isolates the accuracy of the SQD impurity solver relative to FCI within the same embedded active space.

employed 30 qubits, with a symmetry-space dimension of 25,050,025 and a Hilbert-space dimension of 1,073,741,824. The largest quantum circuit depth was 1081. These entries have been highlighted in the table to indicate the maximal values observed in our work. In contrast, the complete (unfragmented) quantum simulation of the molecules with SQD would require the following number of qubits in the STO-3G basis: $HOCN$ (32 qubits), CH_3NO (36 qubits), CH_5NO (40 qubits), C_2H_3NO (46 qubits), C_2H_5NO (50 qubits), CH_4N_2O (48 qubits), $NOCl$ (38 qubits), and $HOSCN$ (50 qubits).

Section E in Supplementary Material provides the detailed orbital space decomposition for each fragment. We note that in DMET, only the impurity space comprising the fragment orbitals A_y and bath orbitals B_y is solved on the quantum computer, requiring $2(|A_y| + |B_y|)$ qubits per fragment at most. The core orbitals Cor_y and virtual orbitals Vir_y are excluded from the quantum simulation entirely - their contributions enter only as a classical mean-field correction to the embedded Hamiltonian. Consequently, there is no additional qubit overhead associated with Cor_y or Vir_y. The largest fragment in this work amounts to a reduction from 50 qubits (full HOSCN) to at most 30 qubits ([S] fragment).

Fig. 5 illustrates the convergence behavior of the DMET-SQD workflow with respect to the number of iterations N_{it} (required for μ_{glob} convergence) for a set of ligand-like molecules. Panel (a) reports the absolute energy difference $\Delta E = |E_{\text{DMET-SQD}} - E_{\text{DMET-FCI}}|$ on a logarithmic scale, while panel (b) shows the corresponding electron number deviation N_e^{err} for each impurity. A systematic and monotonic reduction of errors is observed as the number of iterations increases, with all systems achieving energy deviations below 10^{-5} Ha and electron number errors below 10^{-7} by $N_{it} = 4$. Notably, this convergence behavior is consistent across molecules with varying chemical composition and molecular weight, indicating that the DMET-SQD workflow is robust with respect to system size within the studied regime. These results demonstrate that, despite aggressive fragmentation (one-atom-per-fragment) and minimal basis sets, DMET-SQD is able to recover correlation energies consistent with the reference DMET-FCI solver to within chemical accuracy.

The simulation results showing the energies obtained through DMET-SQD on a quantum computer, DMET-FCI as reference,

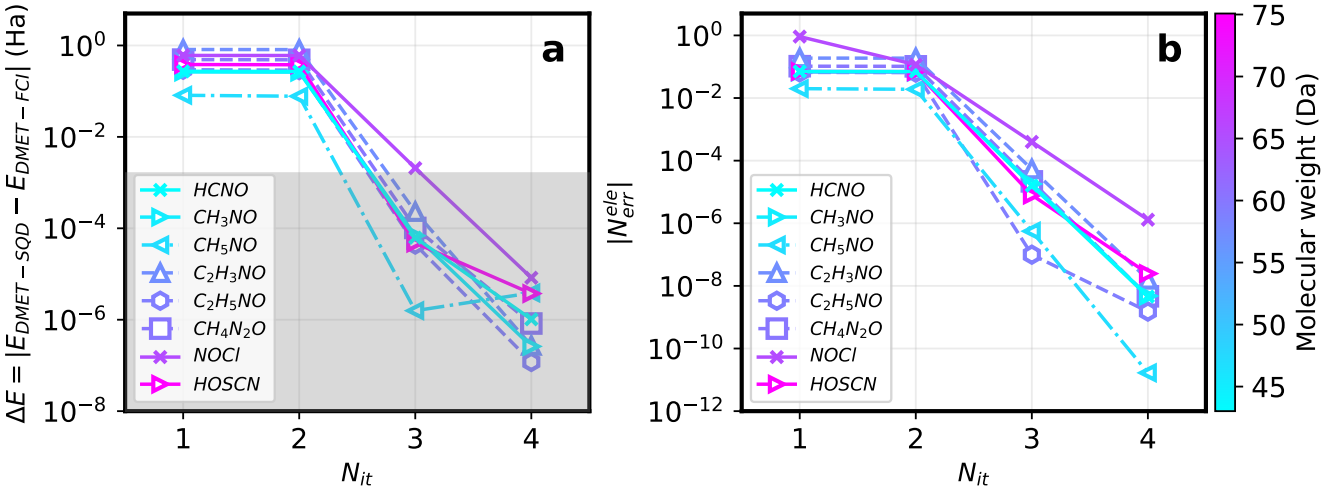

Figure 5. Convergence behavior of the DMET-SQD energies for a set of molecular systems. (a) Absolute energy deviation with respect to the final energy obtained through DMET-FCI, $\Delta E = |E_{\text{DMET-SQD}} - E_{\text{DMET-FCI}}|$, (b) Corresponding absolute electron-number error, $|N_{err}^{ele}| = |\sum_y^{N_{frag}} N_{A_y}^{ele} - N_{true}|$ as given in Eq. (13), as a function of the iteration number N_{it} of the chemical potential μ_{glob} convergence loop. Different molecular systems are distinguished by marker style, while the color scale denotes molecular weight.

and the absolute energy difference between them are shown in Section B in Supplementary Material. The energy differences lie along the range of micro-Hartree, indicating strong agreement with the reference. The geometric co-ordinates of the molecules used for simulation were geometry-optimized (force-field level)[98] and are given in the Section A of the Supplementary Material.

These results indicate that the quantum sampling, configuration recovery, and subsequent creation of projection sub-space through symmetrized expansion of the unique spin configurations (Eq. 9 and Eq. 10), is leading to diagonalization over a sub-space which is a sizable fraction of the *symmetry space*[39].

4.2 Entanglement-aware Embedding via Occupation Threshold ε_{occ}

From a quantum computing perspective, the choice of ε_{occ} plays a critical role in balancing correlation accuracy and hardware feasibility. A smaller cutoff retains more bath orbitals, thereby increasing the impurity size, which in turn raises the number of qubits and circuit depth required for SQD. This leads to a degradation in sampling quality due to increased noise accumulation on NISQ hardware, which requires more robust configuration recovery. Conversely, a larger cutoff reduces the impurity size, improving sampling efficiency and robustness, but at the cost of neglecting weaker correlations. This trade-off is particularly pronounced in low-symmetry (C_1) molecular systems, where geometric details strongly influence fragment-environment entanglement patterns. Our observations indicate that ε_{occ} is therefore a physically and computationally significant control knob that governs both the expressibility of the embedding and the viability of quantum sampling. Notably, this sensitivity is particularly consequential in the DMET-SQD context, where impurity size directly governs qubit count, circuit depth, and shot budget - quantum hardware constraints that are inconsequential for classical impurity solvers such as FCI.

4.2.1 Bath Orbital Construction for various Occupation Thresholds

The construction of bath orbitals in practical DMET implementations relies on the spectrum of the environment block of the 1-RDM, where only a subset of eigenvectors corresponding to nontrivial (fractional) eigenvalues are retained as entangled bath orbitals. In realistic quantum chemistry calculations, several of these eigenvalues may lie numerically very close to 0 or 1, making it difficult to distinguish between genuinely entangled orbitals and effectively unentangled environment orbitals. Therefore, this leads to a practically feasible and efficient construction of the impurity, consisting of the localized fragment orbitals and the entangled bath orbitals with eigenvalues adhering to ε_{occ}, as discussed in the methodology, to truncate negligible contributions and define the bath space[43].

In our DMET-SQD simulations with atom-wise fragmentation, this truncation directly impacts the resulting impurity sizes. For instance, in the cases of NOCl and HOSCN, although the nominal impurity size for the [Cl] and [S] fragments would correspond to 18 orbitals, the application of the eigenvalue cutoff reduces the effective bath to only 3 and 6 orbitals, respectively, yielding smaller impurity spaces consisting of 12 and 15 orbitals, respectively. This phenomenon has been critically analyzed and observed for the HOSCN molecule, as shown in Fig. 6. This reflects that only a limited subset of orbitals significantly

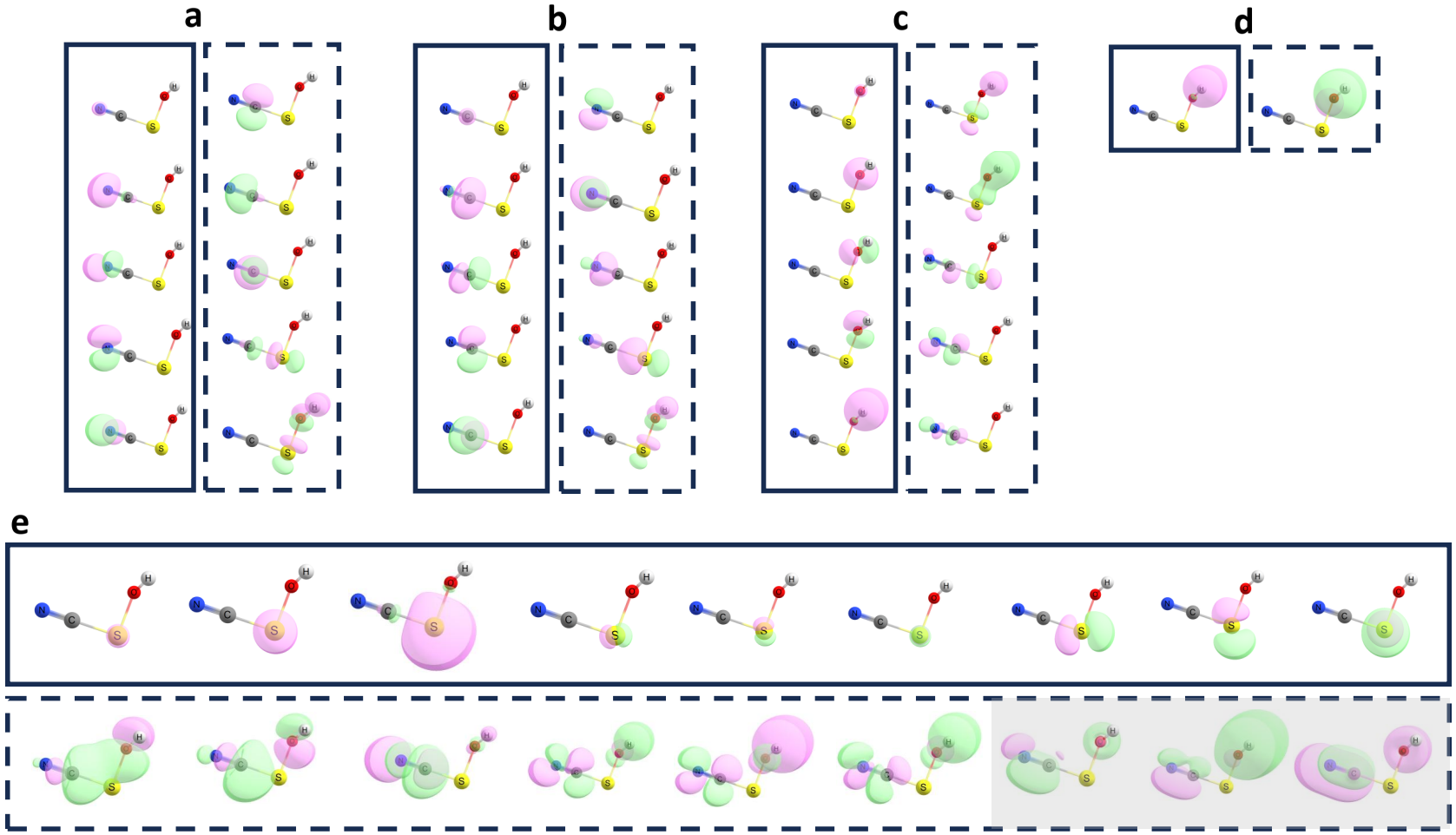

Figure 6. The localized fragment and bath orbitals for the HSOCN molecule for the fragments (a), (b), (c), (d) and (e) denoting the fragments [N], [C], [O], [H], and [S] respectively. The boxes with a solid line denote the localized fragment orbitals, and the boxes with a dashed line denote the corresponding bath orbitals. The last three bath orbitals in subplot (e) are not considered as part of the impurity orbitals based on the value of $\varepsilon_{occ} = 10^{-13}$.

contributes to the fragment–environment entanglement, as determined by the chosen ε_{occ} threshold. This behaviour is consistent with the underlying fragment-environment entanglement structure of these molecules, and its implications for quantum resource requirements and embedding accuracy are examined subsequently.

To quantify the impact of ε_{occ} on the embedding and the resulting quantum resource requirements, we performed a systematic threshold sweep on HOSCN using six values $\varepsilon_{\text{occ}} \in \{10^{-5}, 10^{-7}, 10^{-9}, 10^{-11}, 10^{-13}, 10^{-15}\}$, with all other simulation parameters held fixed. By MacDonald's theorem[86], the rank-$|A_y|$ coupling between the fragment and its environment bounds the number of eigenvalues of the environment block D_y^{env} that can be displaced from 0 or 2, so at most $|A_y|$ fractional (entangled) eigenvalues exist per fragment. The fractionality of each eigenvalue ε of D_y^{env} is measured by $\delta = \min(\varepsilon, 2-\varepsilon)$: for ε close to 0 (empty orbital) or 2 (filled orbital), $\delta \approx 0$ and the orbital is unentangled, whereas $\delta \approx 1$ signals a maximally entangled bath orbital. The MacDonald bound therefore directly limits the number of eigenvalues with $\delta > 0$ to at most $|A_y|$ per fragment. For HOSCN, this yields an upper bound of 1 for [H], 5 for each of [O], [C], and [N], and 9 for [S], consistent with their respective fragment orbital counts $|A_y|$ given in Section E in Supplementary Material.

The environment 1-RDM eigenvalue fractionality spectra per fragment, shown in Fig. 7, directly confirm this bound: each panel contains at most $|A_y|$ eigenvalues with $\delta > 0$, while all remaining environment eigenvalues collapse to the noise floor, corresponding to exactly unentangled core or virtual orbitals. The entanglement structure is, however, highly fragment-dependent in how those $|A_y|$ fractional eigenvalues are distributed: [H], [O], [C], and [N] exhibit sharp spectral gaps, with their few fractional eigenvalues well-separated from the noise floor, so that $|B_y|$ is stable across all thresholds. In contrast, [S] shows a gradual spectrum spanning several orders of magnitude — its 9 fractional eigenvalues decay continuously from $\delta \sim 10^{-1}$ down to $\delta \sim 10^{-14}$ without a clean gap, so that $|B_y|$ grows from 3 to 9 as ε_{occ} is tightened from 10^{-5} to 10^{-15}.

Since ε_{occ} directly controls which of these MacDonald-bounded fractional eigenvalues are admitted as bath orbitals, it serves as a physically grounded, entanglement-aware embedding control parameter: a tighter threshold retains more of the fragment–environment entanglement at the cost of larger impurity spaces, while a looser threshold discards weakly entangled orbitals to reduce quantum resource requirements. This sensitivity is particularly consequential for [S], whose absence of a clean spectral gap makes it uniquely vulnerable to the choice of threshold, with direct consequences for quantum resource requirements and embedding accuracy examined subsequently.

4.2.2 Analysis of ε_{occ} within DMET-SQD

Fig. 8 illustrates the convergence of the DMET-SQD workflow for HOSCN as a function of both N_{it} and the occupation threshold ε_{occ}, revealing that thresholds in the range 10^{-9} to 10^{-13} yield the fastest convergence and the lowest final ΔE, reaching below 10^{-5} Ha by $N_{it} = 4$ while remaining within the chemical accuracy threshold throughout. At the largest threshold ($\varepsilon_{\text{occ}} = 10^{-5}$) and the smallest ($\varepsilon_{\text{occ}} = 10^{-15}$), the final ΔE is significantly larger, as shown in Fig. 8.

Fig. 9 summarizes the quantum circuit resources for HOSCN as a function of ε_{occ}. As expected, the circuit width grows monotonically with decreasing threshold, since a finer ε_{occ} retains more bath orbitals and thereby enlarges the active space. The [S] fragment is the most sensitive to this truncation: owing to its large fragment size ($|A_y| = 9$), it reaches a maximum of 36 qubits at $\varepsilon_{\text{occ}} = 10^{-15}$. Circuit depth follows a similar trend, with [S] exhibiting the deepest circuits across all thresholds.

Fig. F1 in Section F of Supplementary Material presents the classical S-CoRe resource requirements for each DMET fragment of HOSCN as a function of ε_{occ}. For the [H], [O], [C], and [N] fragments, the symmetry-space dimension $|S_{\text{proj}}|$ remains largely stable across all thresholds, with sampling ratios consistently above 95%, indicating that the fixed shot budget of 10^4 is sufficient to cover the projected subspace for these fragments at all thresholds. The [S] fragment presents a qualitatively different picture: $|S_{\text{proj}}|$ grows by nearly two orders of magnitude from $\varepsilon_{\text{occ}} = 10^{-5}$ to 10^{-15}, accompanied by a collapse of the sampling ratio from ∼83% to ∼2%. This reflects the growth of the [S] active space from 26 to 36 qubits as ε_{occ} is tightened, driven by the bath truncation exception $|A_y| \neq |B_y|$ unique to this fragment (see Section E in Supplementary Material). Taken together, these results establish that ε_{occ} functions as an entanglement-aware embedding control parameter: for HOSCN, the threshold 10^{-13} achieves chemical accuracy while avoiding the unnecessary resource overhead incurred by admitting near-zero eigenvalues, providing a concrete and transferable criterion for threshold selection in low-symmetry molecular applications. However, an arbitrary lowering of the ε_{occ} value without increasing the shot budget would lead to under exploration of the Hilbert space, resulting in lower accuracy.

The observations above suggest a practical preliminary analysis protocol for DMET-SQD calculations: prior to any quantum computation, performing the diagonalization of the environment block of the 1-RDM as part of a mean-field DMET-HF calculation and inspecting the resulting fractionality spectrum $\delta = \min(\varepsilon, 2-\varepsilon)$ per fragment provides a zeroth-order diagnostic for the selection of ε_{occ}. The spectral structure (specifically, the presence or absence of a clean gap separating genuinely fractional eigenvalues from the noise floor) directly quantifies the fragment–environment entanglement and anticipates the sensitivity of the impurity construction to the threshold choice. This is computationally inexpensive, requiring only a classical mean-field calculation, yet yields actionable guidance: fragments exhibiting sharp spectral gaps (such as [H], [O], [C], and [N] in HOSCN) are insensitive to ε_{occ} and require no threshold tuning, whereas fragments with a gradual spectrum (such as [S]) signal that the threshold must be chosen with care to balance correlation completeness against quantum and classical

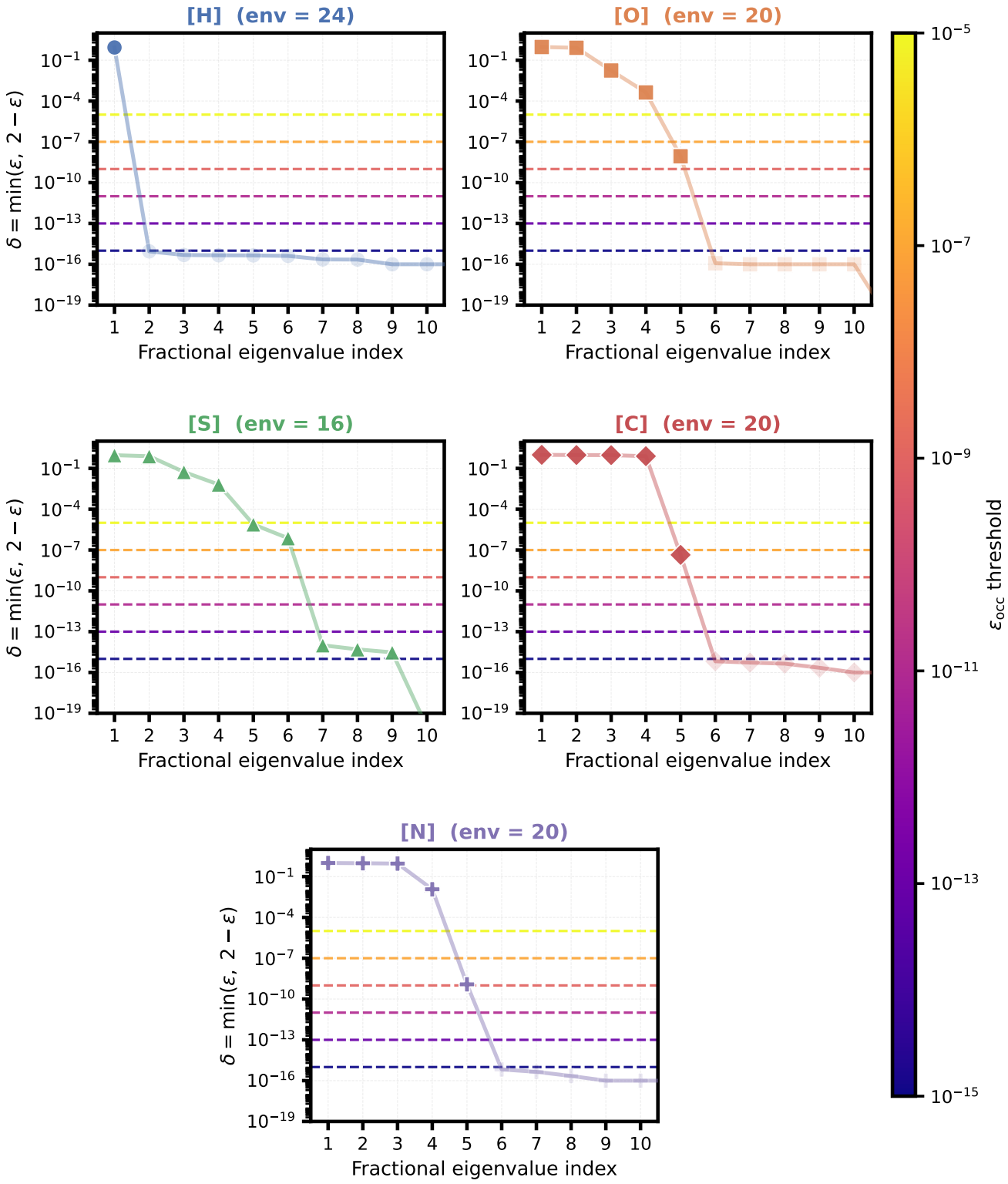

Figure 7. Environment 1-RDM eigenvalue fractionality spectrum $\delta = \min(\varepsilon, 2-\varepsilon)$ for all fragments of HOSCN in the STO-3G basis (**25 total spatial orbitals**), at six occupation thresholds $\varepsilon_{\text{occ}} \in \{10^{-5}, 10^{-7}, 10^{-9}, 10^{-11}, 10^{-13}, 10^{-15}\}$. The environment size per fragment follows directly from $|Env_y| = N_{\text{orb}} - |A_y|$: [H] has $|Env_y| = 24$ ($|A_y| = 1$), [O], [C], and [N] each have $|Env_y| = 20$ ($|A_y| = 5$), and [S] has $|Env_y| = 16$ ($|A_y| = 9$). Dashed horizontal lines mark each threshold, and eigenvectors corresponding to eigenvalues above a given line are selected as bath orbitals at that ε_{occ}.

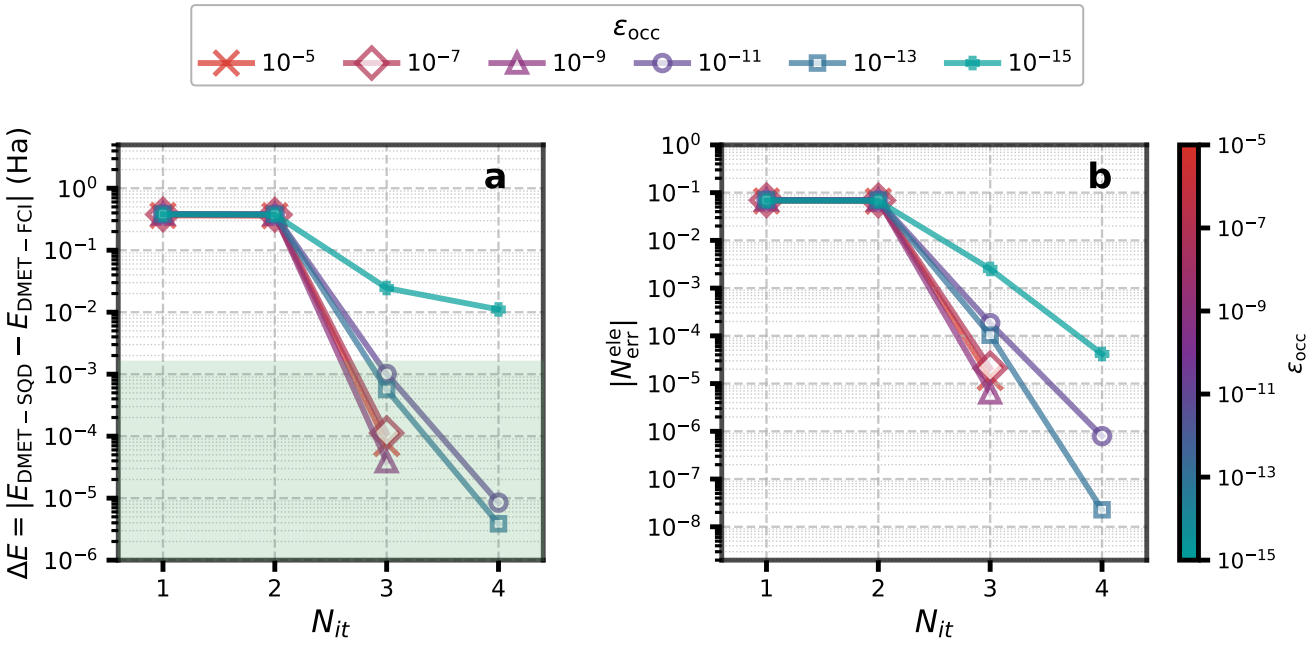

Figure 8. Convergence of the DMET-SQD energy error and global chemical-potential optimization for HOSCN at different occupation thresholds ε_{occ} on IBM Boston Heron R3 (13^{th} June, 2026). (a) Energy error, $\Delta E = |E_{\text{DMET-SQD}} - E_{\text{DMET-FCI}}|$, and (b) electron-number error, $|N_{err}^{ele}|$, as functions of the DMET iteration number N_{it}. Marker styles distinguish different ε_{occ} values, while colours indicate $\log_{10}(\varepsilon_{occ})$, with darker shades corresponding to smaller thresholds. The shaded region in panel (a) denotes the chemical-accuracy regime (1 kcal/mol $\approx 1.594 \times 10^{-3}$ Ha).

resource constraints. In particular, a tighter ε_{occ} admits more bath orbitals and captures more correlation, but enlarges the impurity Hamiltonian whose classical diagonalization and quantum circuit shot budget scale exponentially with active space size, rendering an overly stringent threshold computationally prohibitive.

Taken together, this preliminary spectral analysis elevates ε_{occ} selection from an empirical numerical choice to an entanglement-aware decision: by grounding the threshold in the fractionality spectrum of the mean-field 1-RDM, the entire DMET-SQD workflow (from bath construction and impurity sizing through to quantum sampling and classical diagonalization) is guided by the true fragment-environment entanglement content of the system. Consequently, ε_{occ} is not merely a numerical convergence parameter but a physically transparent, entanglement-aware embedding control parameter whose optimal selection is essential for balancing correlation accuracy against qubit count, circuit depth, and shot budget in low-symmetry molecular systems on quantum hardware. This consideration extends naturally to recent embedding frameworks such as EWF[69] that builds upon DMET as its foundational layer.

Moreover, SQD's reliance on recovering signal from noisy configurations, evidenced by the need for only 2% meaningful signal amid raw noisy data in benchmark cases[28]. This reveals a deep hardware dependence that, while enabling NISQ-scale simulations, introduces inefficiencies tied to noise levels. In lower-noise regimes (higher signal configurations), the subspace dimension could remain fixed, but fewer invalid configurations would require correction via the iterative orbital occupancy updates, potentially shifting the burden to the quantum sampler itself. Sampling from this single ansatz distribution, however, introduces key caveats[31]: notably, the risk of repetitive over-sampling of dominant configurations, leaving scant opportunity to capture rarer determinants that contribute non-negligibly to the true ground state.

Although lower hardware noise increases the fidelity of sampled configurations, it simultaneously suppresses the population of symmetry-breaking configurations. Configuration Recovery relies upon these noisy configurations to expand the sampled subspace using the orbital-occupancy distribution as a reference. In the extreme low-noise limit, this reduced exploration can narrow the recoverable determinant space and ultimately degrade the accuracy of SQD[99]. Modern trapped-ion platforms achieve two-qubit gate fidelities of $\approx 99.92\%$ (Quantinuum)[100], $\approx 99.99\%$ (IONQ)[101], neutral-atom arrays report two-qubit fidelities approaching 99.5% (QuEra)[102], and superconducting qubits have recently achieved $\approx 99.5\%$ (Rigetti)[103], $\approx 99.88\%$ (IBM)[104], $\approx 99.93\%$ (IQM)[105] two-qubit fidelities, all of which influence the balance between valid and recoverable configurations in SQD.

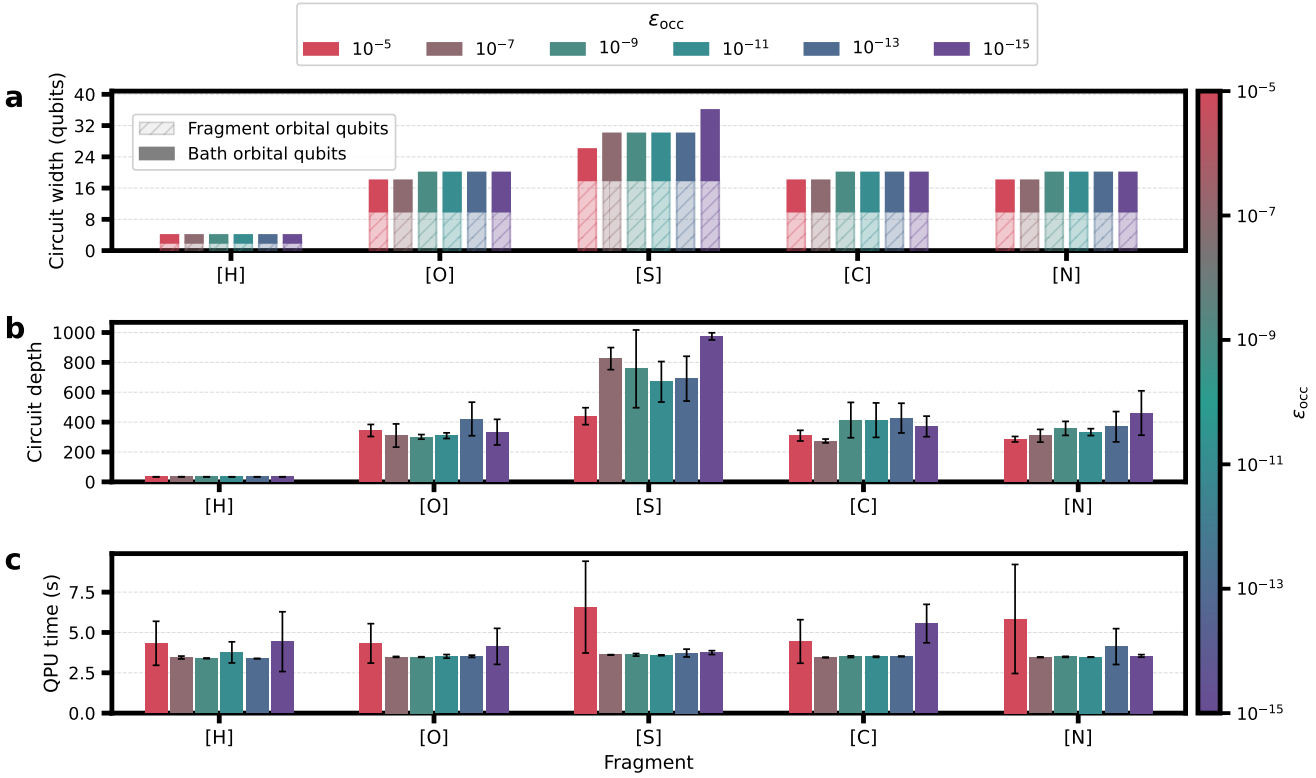

Figure 9. Quantum circuit resources for HOSCN/STO-3G across ε_{occ} thresholds Circuit width decomposed into fragment orbital qubits ($2|A_y|$, hatched) and bath orbital qubits ($2|B_y|$, solid) (a), transpiled circuit depth (b), and QPU execution time (c) for each DMET fragment of HOSCN, executed on the IBM Heron R3 (`ibm_boston`) quantum processor across six thresholds $\varepsilon_{\text{occ}} \in \{10^{-5}, 10^{-7}, 10^{-9}, 10^{-11}, 10^{-13}, 10^{-15}\}$. Error bars denote $\pm 1\sigma$ over all μ_{glob} convergence iterations.

5 Conclusion

In this work, we have successfully performed quantum simulations of a set of natural ligand-like molecules in the STO-3G basis set (with molecular weights between 40 Da and 76 Da). We employed DMET as a classical fragmentation technique followed by SQD as a high-level solver. The simulations were run on IBM's Eagle R3 (`ibm_sherbrooke`) superconducting quantum computing hardware. All the results obtained were benchmarked against DMET-FCI and were found to be within chemical accuracy.

In addition, a detailed entanglement-aware analysis of the sample-based technique within the embedding framework was performed for the HOSCN molecule by systematically varying ε_{occ} across six orders of magnitude on IBM's Heron R3 (`ibm_boston`). This analysis demonstrated that ε_{occ} functions as a physically meaningful, entanglement-aware control parameter rather than a simple numerical convergence threshold, with a fragment-dependent sensitivity that is particularly pronounced for strongly entangled fragments such as [S], and obtaining an optimal threshold range might be beneficial, which balances embedding completeness against quantum hardware feasibility.

Thereby, the current NISQ hardware seems to show potential to perform quantum sampling experiments, leading to classical sub-space diagonalization in post-processing with high accuracy. This paves the way for potential industrial use-cases in the field of quantum computer-aided drug design and materials design. In practice, an overly stringent threshold increases circuit depth and noise sensitivity without a comparable accuracy benefit, while an overly relaxed threshold risks neglecting non-negligible fragment-environment correlation. For HOSCN, the environment 1-RDM eigenvalue spectrum (Fig. 7) provided a direct, per-fragment diagnostic for identifying this balance; extending such a diagnostic-driven selection of ε_{occ} to other low-symmetry systems remains an important direction for future work.

Notably, this trade-off is not merely qualitative: our threshold sweep on HOSCN shows that the energy accuracy ΔE depends non-monotonically on ε_{occ}, with the tightest threshold ($\varepsilon_{occ} = 10^{-15}$) yielding *worse* accuracy than an intermediate value ($\varepsilon_{occ} = 10^{-13}$) due to severe under-sampling of the enlarged [S] impurity. This finding cautions against the common assumption that tighter occupation thresholds are uniformly preferable, and instead establishes ε_{occ} selection as an explicit accuracy–resource optimization problem for low-symmetry molecular systems on near-term hardware.

That being said, although SQD ensures that the noisy configurations are recovered into the symmetry space by using the average orbital occupancy distributions as reference, this does not guarantee that the noisy configurations are stochastically converted into determinants with a non-negligible contribution to the true ground state. The procedure still might lead to obtaining configurations from the symmetry space, which contribute little to the final energy estimate.

This highlights the need for approaches that explicitly promote compactness in the selected subspace, where the retained configurations are guided by user-defined criteria rather than by sampling alone. By constraining the search to a compact and size-controlled subspace, one can more reliably concentrate quantum and classical resources only on the determinants most relevant to the target state, leading to an overall reduction of computational cost for the procedure.

This need is further underscored by the fact that the performance of S-CoRe is intrinsically tied to the noise level of the hardware: in ultra-low-noise regimes, the sampling distribution collapses onto a few high-weight determinants concentrated near the Hartree-Fock reference, severely limiting subspace diversity, whereas moderate noise increases the proportion of off-symmetry samples that, after configuration-recovery, can populate a richer and more correlation-relevant subspace. Relying solely on stochastic sampling and symmetry recovery, therefore, leaves subspace quality coupled to hardware noise characteristics outside the user's control, motivating more explicitly guided subspace construction strategies[39–42].

A further promising direction is the incorporation of cumulant-based 2-RDM reconstruction[106] within the DMET-SQD framework, which has been shown to improve the accuracy of global observables and is in principle compatible with the RDM outputs of the SQD procedure. Additionally, extension of the present framework to larger basis sets and open-shell systems, as well as the systematic study of the effect of ε_{occ} across the full molecular set studied here, remain important avenues for future investigation.

Supplemetary Material

The Supplementary Material provides comprehensive supporting data for the present study, including the optimized geometric coordinates of all ligand-like molecules considered, detailed comparisons of ground-state energies obtained using DMET-SQD and DMET-FCI (demonstrating good agreement with chemical accuracy), the calibration parameters of both the IBM Eagle R3 (`ibm_sherbrooke`) and IBM Heron R3 (`ibm_boston`) quantum processors employed in the simulations, the complete fragment orbital space decomposition for all studied molecules, and the classical reosurce cost for HOSCN molecule while varying the occupancy threshold.

Acknowledgements

We acknowledge the use of IBM Quantum Credits via the IBM Quantum Startups Program for this work. The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Quantum Platform team. The authors would also like to extend appreciation to the advisors of Qclairvoyance Quantum Labs for their support, constructive discussions, and inspiration throughout the preparation of this work.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests

R.M. and R.V. are paid consultants at Qclairvoyance Quantum Labs. The other authors declare no competing interests.

References

1. Levine, I. N. *Molecular Orbital Theory*, chap. 16, 653–710 (Pearson Education, Boston, 2014), 7 edn.
2. Shayit, A. *et al.* Numerically exact configuration interaction at quadrillion-determinant scale. *Nat. Commun.* **16**, 11016 (2025), DOI: 10.1038/s41467-025-65967-7.
3. Franzke, Y. J., Treß, R., Pazdera, T. M. & Weigend, F. Error-consistent segmented contracted all-electron relativistic basis sets of double- and triple-zeta quality for nmr shielding constants. *Phys. Chem. Chem. Phys.* **21**, 16658–16664 (2019), DOI: 10.1039/C9CP02382H.
4. Gao, H., Imamura, S., Kasagi, A. & Yoshida, E. Distributed implementation of full configuration interaction for one trillion determinants. *J. Chem. Theory Comput.* **20**, 1185–1192 (2024), DOI: 10.1021/acs.jctc.3c01190.
5. Hehre, W. J., Stewart, R. F. & Pople, J. A. Self-consistent molecular-orbital methods. i. use of gaussian expansions of slater-type atomic orbitals. *The J. Chem. Phys.* **51**, 2657–2664 (1969), DOI: 10.1063/1.1672392.

6. Fock, V. Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems. *Zeitschrift für Physik* **61**, 126–148 (1930), DOI: 10.1007/BF01340294.

7. Kirsopp, J. J. M. *et al.* Quantum computational quantification of protein–ligand interactions. *Int. J. Quantum Chem.* **122**, e26975 (2022), DOI: 10.1002/qua.26975.

8. Malone, F. D. *et al.* Towards the simulation of large scale protein–ligand interactions on nisq-era quantum computers. *Chem. Sci.* **13**, 3094–3108 (2022), DOI: 10.1039/D1SC05691C.

9. Bowling, P. E., Broderick, D. R. & Herbert, J. M. Convergent protocols for computing protein–ligand interaction energies using fragment-based quantum chemistry. *J. Chem. Theory Comput.* **21**, 951–966 (2025), DOI: 10.1021/acs.jctc.4c01429.

10. Barton, D. H. R. & Cookson, R. C. The principles of conformational analysis. *Q. Rev. Chem. Soc.* **10**, 44–82 (1956), DOI: 10.1039/QR9561000044.

11. Mancini, G., Fusè, M., Lazzari, F. & Barone, V. Fast exploration of potential energy surfaces with a joint venture of quantum chemistry, evolutionary algorithms and unsupervised learning. *Digit. Discov.* **1**, 790–805 (2022), DOI: 10.1039/D2DD00070A.

12. Anurag, K. S. V. *et al.* Potential energy surface scan of *n*-butane using various quantum chemistry software. *Authorea* (2025), DOI: 10.22541/au.174624847.78353339/v2.

13. Hao, Y., Ding, Q., Wang, X. & Yuan, X. Large-scale efficient molecule geometry optimization with hybrid quantum-classical computing. *arXiv* (2025), DOI: 10.48550/arXiv.2509.07460.

14. Shajan, A. *et al.* Toward quantum-centric simulations of extended molecules: Sample-based quantum diagonalization enhanced with density matrix embedding theory. *J. Chem. Theory Comput.* **21**, 6801–6810 (2025), DOI: 10.1021/acs.jctc.5c00114.

15. Kawashima, Y. *et al.* Optimizing electronic structure simulations on a trapped-ion quantum computer using problem decomposition. *Commun. Phys.* **4**, 245 (2021), DOI: 10.1038/s42005-021-00751-9.

16. O'Malley, P. J. J. *et al.* Scalable quantum simulation of molecular energies. *Phys. Rev. X* **6**, 031007 (2016), DOI: 10.1103/PhysRevX.6.031007.

17. Schuch, N. & Verstraete, F. Computational complexity of interacting electrons and fundamental limitations of density functional theory. *Nat. Phys.* **5**, 732–735 (2009), DOI: 10.1038/nphys1370.

18. Cao, Y. *et al.* Quantum chemistry in the age of quantum computing. *Chem. Rev.* **119**, 10856–10915 (2019), DOI: 10.1021/acs.chemrev.8b00803.

19. Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **309**, 1704–1707 (2005), DOI: 10.1126/science.1113479.

20. Eisert, J. & Preskill, J. Mind the gaps: The fraught road to quantum advantage. *arXiv* **arXiv:2510.19928** (2025), DOI: 10.48550/arXiv.2510.19928.

21. Kitaev, A. Y. Quantum measurements and the abelian stabilizer problem. *arXiv:quant-ph/9511026* (1995).

22. Gilyén, A., Su, Y., Low, G. H. & Wiebe, N. Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* STOC '19, 193–204, DOI: 10.1145/3313276.3316366.

23. Wang, X. *et al.* Randomized quantum singular value transformation. *arXiv* (2025), DOI: 10.48550/arXiv.2510.06851.

24. Preskill, J. Quantum computing in the nisq era and beyond. *Quantum* **2**, 79 (2018), DOI: 10.22331/q-2018-08-06-79.

25. Peruzzo, A. *et al.* A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014), DOI: 10.1038/ncomms5213.

26. Belaloui, N. E. *et al.* Ground-state energy estimation on current quantum hardware through the variational quantum eigensolver: A practical study. *J. Chem. Theory Comput.* **21**, 6777–6792 (2025), DOI: 10.1021/acs.jctc.4c01657.

27. Tilly, J. *et al.* The variational quantum eigensolver: A review of methods and best practices. *Phys. Reports* **986**, 1–128 (2022), DOI: 10.1016/j.physrep.2022.08.003.

28. Robledo-Moreno, J. *et al.* Chemistry beyond the scale of exact diagonalization on a quantum-centric supercomputer. *Sci. Adv.* **11**, eadu9991 (2025), DOI: 10.1126/sciadv.adu9991.

29. Kanno, K. *et al.* Quantum-selected configuration interaction: Classical diagonalization of hamiltonians in subspaces selected by quantum computers. *arXiv* (2023), DOI: 10.48550/arXiv.2302.11320.

30. Evangelista, F. A. Adaptive multiconfigurational wave functions. *The J. Chem. Phys.* **140**, 124114 (2014), DOI: 10.1063/1.4869192.

31. Reinholdt, P. *et al.* Critical limitations in quantum-selected configuration interaction methods. *J. Chem. Theory Comput.* **21**, 6811–6822 (2025), DOI: 10.1021/acs.jctc.5c00375.

32. Alexeev, Y. *et al.* Quantum-centric supercomputing for materials science: A perspective on challenges and future directions. *Futur. Gener. Comput. Syst.* **160**, 666–710 (2024), DOI: 10.1016/j.future.2024.04.060.

33. Barison, S., Robledo Moreno, J. & Motta, M. Quantum-centric computation of molecular excited states with extended sample-based quantum diagonalization. *Quantum Sci. Technol.* **10**, 025034 (2025), DOI: 10.1088/2058-9565/adb781.

34. Yoshioka, N. *et al.* Krylov diagonalization of large many-body hamiltonians on a quantum processor. *Nat. Commun.* **16**, 5014 (2025), DOI: 10.1038/s41467-025-59716-z.

35. Campbell, E. Random compiler for fast hamiltonian simulation. *Phys. Rev. Lett.* **123**, 070503 (2019), DOI: 10.1103/PhysRevLett.123.070503.

36. Yu, J. *et al.* Quantum-centric algorithm for sample-based krylov diagonalization. *arXiv* (2025), DOI: 10.48550/arXiv.2501.09702.

37. Asthana, A. Quantum krylov algorithm using unitary decomposition for exact eigenstates of fermionic systems using quantum computers. *arXiv* (2025), DOI: 10.48550/arXiv.2512.11788.

38. Piccinelli, S. *et al.* Quantum chemistry with provable convergence via randomized sample-based quantum diagonalization. *arXiv* (2025), DOI: 10.48550/arXiv.2508.02578.

39. Pellow-Jarman, A. *et al.* Hivqe: Handover iterative variational quantum eigensolver for efficient quantum chemistry calculations. *arXiv* (2025), DOI: 10.48550/arXiv.2503.06292.

40. Yoo, P. *et al.* Extending the handover-iterative vqe to challenging strongly correlated systems: n_2 and fe-s cluster. *arXiv* (2026), DOI: 10.48550/arXiv.2601.06935.

41. Patra, C. *et al.* Physics-informed generative machine learning for accelerated quantum-centric supercomputing. *arXiv* (2026), DOI: 10.48550/arXiv.2512.06858.

42. Patra, A. K. *et al.* Machine-learned compact subspace generation for quantum selected configuration interaction within density matrix embedding framework (2026). Manuscript in preparation.

43. Wouters, S., Jiménez-Hoyos, C. A., Sun, Q. & Chan, G. K.-L. A practical guide to density matrix embedding theory in quantum chemistry. *J. Chem. Theory Comput.* **12**, 2706–2719 (2016), DOI: 10.1021/acs.jctc.6b00316.

44. Knizia, G. & Chan, G. K.-L. Density matrix embedding: A simple alternative to dynamical mean-field theory. *Phys. Rev. Lett.* **109**, 186404 (2012), DOI: 10.1103/PhysRevLett.109.186404.

45. Negre, A. *et al.* New perspectives on density-matrix embedding theory. *arXiv* (2025), DOI: 10.48550/arXiv.2503.09881.

46. Cancès, E., Faulstich, F. M., Kirsch, A., Letournel, E. & Levitt, A. Analysis of density matrix embedding theory around the non-interacting limit. *Commun. on Pure Appl. Math.* **78**, 1359–1410 (2025), DOI: https://doi.org/10.1002/cpa.22244. https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.22244.

47. Pham, H. Q., Bernales, V. & Gagliardi, L. Can density matrix embedding theory with the complete active space self-consistent field solver describe single and double bond breaking in molecular systems? *J. Chem. Theory Comput.* **14**, 1960–1968 (2018), DOI: 10.1021/acs.jctc.7b01248.

48. Verma, S., Hermes, M. R. & Gagliardi, L. Density matrix embedding pair-density functional theory for molecules. *The J. Phys. Chem. Lett.* **16**, 5348–5357 (2025), DOI: 10.1021/acs.jpclett.5c00829.

49. Nakai, H. *et al.* Divide-and-conquer linear-scaling quantum chemical computations. *The J. Phys. Chem. A* **127**, 589–618 (2023), DOI: 10.1021/acs.jpca.2c06965.

50. Araujo, I. F., Park, D. K., Petruccione, F. & da Silva, A. J. A divide-and-conquer algorithm for quantum state preparation. *Sci. Reports* **11**, 6329 (2021), DOI: 10.1038/s41598-021-85474-1.

51. Fedorov, D. G., Nagata, T. & Kitaura, K. Exploring chemistry with the fragment molecular orbital method. *Phys. Chem. Chem. Phys.* **14**, 7562–7577 (2012), DOI: 10.1039/C2CP23784A.

52. Yamazaki, T., Matsuura, S., Narimani, A., Saidmuradov, A. & Zaribafiyan, A. Towards the practical application of near-term quantum computers in quantum chemistry simulations: A problem decomposition approach. *arXiv* (2018), DOI: 10.48550/arXiv.1806.01305.

53. Sun, Q. & Chan, G. K.-L. Quantum embedding theories. *Accounts Chem. Res.* **49**, 2705–2712 (2016), DOI: 10.1021/acs.accounts.6b00356.

54. Bauman, N. P. & Kowalski, K. Coupled cluster downfolding methods: The effect of double commutator terms on the accuracy of ground-state energies. *The J. Chem. Phys.* **156**, 094106 (2022), DOI: 10.1063/5.0076260.

55. Bauman, N. P. *et al.* Coupled cluster downfolding theory in simulations of chemical systems on quantum hardware. *arXiv* (2025), DOI: 10.48550/arXiv.2507.01199.

56. Hermes, M. R., Pandharkar, R. & Gagliardi, L. Variational localized active space self-consistent field method. *J. Chem. Theory Comput.* **16**, 4923–4937 (2020), DOI: 10.1021/acs.jctc.0c00222.

57. Otten, M. *et al.* Localized quantum chemistry on quantum computers. *J. Chem. Theory Comput.* **18**, 7205–7217 (2022), DOI: 10.1021/acs.jctc.2c00388.

58. Nusspickel, M. & Booth, G. H. Systematic improvability in quantum embedding for real materials. *Phys. Rev. X* **12**, 011046 (2022), DOI: 10.1103/PhysRevX.12.011046.

59. Ma, H. *et al.* Multiscale quantum algorithms for quantum chemistry. *Chem. Sci.* **14**, 3190–3205 (2023), DOI: 10.1039/D2SC06875C.

60. Xu, E., Shimomoto, Y., Ten-no, S. L. & Tsuchimochi, T. Many-body-expansion based on variational quantum eigensolver and deflation for dynamical correlation. *The J. Phys. Chem. A* **128**, 2507–2521 (2024), DOI: 10.1021/acs.jpca.4c00351.

61. Kowalski, K. & Bauman, N. P. Resource-adaptive quantum flow algorithms for quantum simulations of many-body systems: sub-flow embedding procedures. *arXiv* (2024), DOI: 10.48550/arXiv.2410.11992.

62. Georges, A., Kotliar, G., Krauth, W. & Rozenberg, M. J. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.* **68**, 13–125 (1996), DOI: 10.1103/RevModPhys.68.13.

63. Blumenthal, E. Building intuition for dynamical mean-field theory: A simple model and the cavity method. *arXiv* (2025), DOI: 10.48550/arXiv.2507.16654.

64. Ye, H.-Z., Tran, H. K. & Van Voorhis, T. Bootstrap embedding for large molecular systems. *J. Chem. Theory Comput.* **16**, 5035–5046 (2020), DOI: 10.1021/acs.jctc.0c00438.

65. Liu, Y. *et al.* Bootstrap embedding on a quantum computer. *J. Chem. Theory Comput.* **19**, 2230–2247 (2023), DOI: 10.1021/acs.jctc.3c00012.

66. Hardikar, T. *et al.* Quanta-bind: A quantum computing pipeline for modeling strongly correlated metal-protein interactions. In *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)* 538–544 (2024), DOI: 10.1109/QCE60285.2024.00069.

67. Cho, M. *et al.* Quemb: A toolbox for bootstrap embedding calculations of molecular and periodic systems. *J. Phys. Chem. A* **129**, 6538–6551 (2025), DOI: 10.1021/acs.jpca.5c02983.

68. Iijima, N. *et al.* Towards accurate quantum chemical calculations on noisy quantum computers (2023). 2311.09634.

69. Shajan, A. *et al.* Molecular quantum computations on a protein. *arXiv* (2026), DOI: 10.48550/arXiv.2512.17130.

70. Fujii, K. *et al.* Deep variational quantum eigensolver: A divide-and-conquer method for solving a larger problem with smaller size quantum computers. *PRX Quantum* **3**, 010346 (2022), DOI: 10.1103/PRXQuantum.3.010346.

71. Lim, H. *et al.* Fragment molecular orbital-based variational quantum eigensolver for quantum chemistry in the age of quantum computing. *Sci. Reports* **14**, 2422 (2024), DOI: 10.1038/s41598-024-52926-3.

72. Bierman, J. & Liu, Y. Towards utility-scale electronic structure with sample-based quantum bootstrap embedding. *Digit. Discov.* – (2026), DOI: 10.1039/D5DD00416K.

73. Caimi, P. F. *et al.* Phase i clinical trial of the base excision repair inhibitor methoxyamine in combination with fludarabine for patients with advanced hematologic malignancies. *Oncotarget* **8**, 79864–79875 (2017), DOI: 10.18632/oncotarget.20094.

74. Eads, J. R. *et al.* Phase i clinical trial of temozolomide and methoxyamine (trc-102), an inhibitor of base excision repair, in patients with advanced solid tumors. *Investig. New Drugs* **39**, 142–151 (2021), DOI: 10.1007/s10637-020-00962-x.

75. Ghosh, A. K. & Brindisi, M. Urea derivatives in modern drug discovery and medicinal chemistry. *J. Medicinal Chem.* **63**, 2751–2788 (2020), DOI: 10.1021/acs.jmedchem.9b01541.

76. Listro, R. *et al.* Urea-based anticancer agents: Exploring 100 years of research with an eye to the future. *Front. Chem.* **10** (2022), DOI: 10.3389/fchem.2022.995351.

77. Born, M. & Oppenheimer, R. Zur Quantentheorie der Molekeln. *Annalen der Physik* **389**, 457–484 (1927), DOI: 10.1002/andp.19273892002.

78. Szabo, A. & Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Books on Chemistry (Dover Publications, 1996). See Section 2.4.1, pp. 89–95.

79. Jordan, P. & Wigner, E. Über das paulische äquivalenzverbot. *Zeitschrift für Physik* **47**, 631–651 (1928), DOI: 10.1007/BF01331938.

80. 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. *Chem. Sci.* **14**, 11213–11227 (2023), DOI: 10.1039/D3SC02516K.

81. Bartlett, R. J. & Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **79**, 291–352 (2007), DOI: 10.1103/RevModPhys.79.291.

82. Jensen, F. *Introduction to Computational Chemistry* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2006). Second Edition, see Section 4.9, pp. 169–174.

83. Matsuzawa, Y. & Kurashige, Y. Jastrow-type decomposition in quantum chemistry for low-depth quantum circuits. *J. Chem. Theory Comput.* **16**, 944–952 (2020), DOI: 10.1021/acs.jctc.9b00963.

84. AbuGhanem, M. Ibm quantum computers: Evolution, performance, and future directions. *The J. Supercomput.* **81**, 687–715 (2025), DOI: 10.1007/s11227-025-07047-7.

85. Horn, R. & Johnson, C. *Matrix Analysis*, chap. 4.3, 242–248 (Cambridge University Press, 2012), 2nd edn.

86. MacDonald, J. K. L. Successive approximations by the rayleigh–ritz variation method. *Phys. Rev.* **43**, 830–833 (1933), DOI: 10.1103/PhysRev.43.830.

87. Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **17**, 87–94 (1975), DOI: 10.1016/0021-9991(75)90065-0.

88. Sun, Q. *et al.* The python-based simulations of chemistry framework (pyscf). *arXiv* (2017), DOI: 10.48550/arXiv.1701.08223.

89. Papakonstantinou, J. M. & Tapia, R. A. Origin and evolution of the secant method in one dimension. *The Am. Math. Mon.* **120**, 500–518 (2013), DOI: 10.4169/amer.math.monthly.120.06.500.

90. Senicourt, V. *et al.* Tangelo: An open-source python package for end-to-end chemistry workflows on quantum computers. *arXiv* (2022), DOI: 10.48550/arXiv.2206.12424.

91. The ffsim developers. ffsim: Faster simulations of fermionic quantum circuits, v0.0.56 (2025). Accessed: 2025-11-11.

92. Javadi-Abhari, A. *et al.* Quantum computing with qiskit. *arXiv* (2024), DOI: 10.48550/arXiv.2405.08810.

93. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **17**, 261–272 (2020), DOI: 10.1038/s41592-019-0686-2.

94. Ai, Y., Sun, Q. & Jiang, H. Efficient multiconfigurational quantum chemistry approach to single-ion magnets based on density matrix embedding theory. *The J. Phys. Chem. Lett.* **13**, 10627–10634 (2022), DOI: 10.1021/acs.jpclett.2c02890.

95. Mitra, A., Pham, H. Q., Pandharkar, R., Hermes, M. R. & Gagliardi, L. Excited states of crystalline point defects with multireference density matrix embedding theory. *The J. Phys. Chem. Lett.* **12**, 11688–11694 (2021), DOI: 10.1021/acs.jpclett.1c03229.

96. Ai, Y., Li, Z.-W., Guan, Z.-B. & Jiang, H. Density matrix embedding theory-based multiconfigurational quantum chemistry approach to lanthanide single-ion magnets. *J. Chem. Theory Comput.* **21**, 9631–9640 (2025), DOI: 10.1021/acs.jctc.5c01336.

97. Li, G., Ding, Y. & Xie, Y. Tackling the qubit mapping problem for nisq-era quantum devices. *arXiv* (2019), DOI: 10.48550/arXiv.1809.02573.

98. Landrum, G. *et al.* Rdkit: Open-source cheminformatics software(2025), DOI: 10.5281/zenodo.17232453.

99. Anurag, K. S. V. *et al.* Bridging the nisq and fault-tolerant regimes: Generative-ml-assisted quantum selected ci for molecular simulations (2026). Manuscript in preparation.

100. Ransford, A. *et al.* Helios: A 98-qubit trapped-ion quantum computer. *arXiv* (2025), DOI: 10.48550/arXiv.2511.05465.

101. Hughes, A. C. *et al.* Trapped-ion two-qubit gates with >99.99% fidelity without ground-state cooling. *arXiv* (2025), DOI: 10.48550/arXiv.2510.17286.

102. Evered, S. J. *et al.* High-fidelity parallel entangling gates on a neutral-atom quantum computer. *Nature* **622**, 268–272 (2023), DOI: 10.1038/s41586-023-06481-y.

103. Du, Z. *et al.* Optimizing inter-chip coupler link placement for modular and chiplet quantum systems. *arXiv* (2025), DOI: 10.48550/arXiv.2509.10409.

104. IBM Quantum. Ibm quantum — heron & ibm boston processor listings. https://quantum.cloud.ibm.com/computers?processorType=Heron&system=ibm_boston (2026). Accessed: 28 January 2026.

105. Marxer, F. *et al.* Above 99.9% fidelity single-qubit gates, two-qubit gates, and readout in a single superconducting quantum device. *arXiv* (2025), DOI: 10.48550/arXiv.2508.16437.

106. Nusspickel, M. & Booth, G. H. Systematic improvability in quantum embedding for real materials. *Phys. Rev. X* **12**, 011046 (2022), DOI: 10.1103/PhysRevX.12.011046.

107. Krantz, P. *et al.* A quantum engineer's guide to superconducting qubits. *Appl. Phys. Rev.* **6**, 021318 (2019), DOI: 10.1063/1.5089550.

SUPPLEMENTARY MATERIAL

A Geometric Coordinates of the Studied Molecules

Molecule	Geometric Coordinates			
	Atom	x	y	z
Cyanic Acid ($HOCN$)	H	-1.458587	-0.272810	0.065495
	O	-0.600578	0.099568	-0.321249
	C	0.544407	0.090463	0.422709
	N	1.514757	0.082779	1.053115
Formaldehyde Oxime (CH_3NO)	C	-0.910484	-0.020915	-0.332484
	H	2.170747	-0.116917	0.612439
	H	-1.005526	1.043453	-0.507441
	H	-1.760983	-0.667161	-0.500891
	N	0.207852	-0.523560	0.074740
	O	1.298394	0.285100	0.293422
Methoxyamine (CH_5NO)	C	-0.966149	0.201296	0.149956
	H	-1.000265	0.965450	-0.658496
	H	-1.667205	0.516700	0.949826
	H	-1.306933	-0.785003	-0.237043
	H	1.882536	0.441132	-0.409602
	H	1.601346	-1.163838	-0.023188
	N	1.132081	-0.281427	-0.328214
	O	0.324588	0.105691	0.692069
Methyl Isocyanate (C_2H_3NO)	C	-0.872655	-0.048579	-0.012732
	C	1.498758	0.200579	-0.121172
	H	-0.658520	-0.731912	0.839315
	H	-1.419752	-0.612203	-0.796487
	H	-1.514077	0.782782	0.345983
	N	0.351223	0.500017	-0.581535
	O	2.615024	-0.090684	0.326629
Acetaldehyde Oxime (C_2H_5NO)	C	-1.321954	0.079847	-0.059915
	C	0.125140	0.194559	0.278307
	H	-1.848828	-0.492848	0.731931
	H	-1.772167	1.091808	-0.134570
	H	-1.440397	-0.441345	-1.033016
	H	0.567039	1.175157	0.409184
	H	2.659033	0.076802	0.861245
	N	0.845312	-0.873562	0.413794
	O	2.186822	-0.810420	0.726078
Carbamide / Urea (CH_4N_2O)	C	0.023609	0.174787	0.475047
	H	2.152264	-0.041841	0.244225
	H	1.187776	-0.518005	-1.199231
	H	-2.122930	0.259006	0.346006
	H	-1.304030	-0.342648	-1.139910
	N	-1.243207	0.016569	-0.161562
	N	1.224229	-0.157069	-0.220304
	O	0.082288	0.609201	1.655728
Nitrosyl Chloride ($NOCl$)	N	-0.222631	0.458624	0.000000
	O	-1.157638	-0.281711	0.000000
	Cl	1.380269	-0.176914	0.000000
Hydroxythiocyanate ($HOSCN$)	H	-1.576204	0.567554	-0.304395
	O	-1.272167	0.443194	0.630206
	S	-0.421261	-1.048230	0.616500
	C	1.126988	-0.244750	0.353656
	N	2.142645	0.282231	0.181157

Table A1. Geometric co-ordinates of the molecules used for simulation(geometry-optimized upto force-field level)[98].

B Energy Results

Molecule	$E_{\text{DMET-SQD}}$ **(Ha)**	$E_{\text{DMET-FCI}}$ **(Ha)**
Cyanic Acid ($HOCN$)	-165.80748733	-165.80748835
Formaldehyde Oxime (CH_3NO)	-166.89443901	-166.89443875
Methoxyamine (CH_5NO)	-168.04584566	-168.04584180
Methyl Isocyanate (C_2H_3NO)	-204.57166716	-204.57166689
Acetaldehyde Oxime (C_2H_5NO)	-205.56816508	-205.56816496
Carbamide / Urea (CH_4N_2O)	-221.44130237	-221.44130142
Nitrosyl Chloride ($NOCl$)	-582.38458086	-582.38458236
Hydroxythiocyanate ($HOSCN$)	-559.05055857	-559.05056231

Table B1. Energy results obtained from the DMET-SQD simulations for the studied molecules along with the energy reference $E_{\text{DMET-FCI}}$ in Hartree (Ha).

C Quantum Computing Hardware Calibration Data - IBM Sherbrooke

Parameter	**Value**
T1	270.447138 μs
T2	211.026521 μs
Frequency	4.794027 GHz
Anharmonicity	-0.310939 GHz
Readout assignment error	0.021240
P(0 | 1)	0.020507
P(1 | 0)	0.019042
Readout length	1216.0 ns
Identity gate error	0.000244
R_z gate error	0.0
$\sqrt{X}$ (S_x) gate error	0.000244
Pauli-X gate error	0.000244
ECR gate error	0.006767
Gate time	533.333333 ns

Table C1. Calibration metrics for the Eagle R3 processor (IBM Sherbrooke, containing 127 qubits) recorded on 4 June 2025 at 01:48:22 EST. These values correspond to the device's pre-execution calibration snapshot, including coherence times (T_1, T_2), qubit frequency, anharmonicity, readout assignment errors, gate error rates, and gate durations[107]. Please note that the quantum hardware is usually calibrated every 24 hours.

D Quantum Computing Hardware Calibration Data - IBM Boston

Parameter	Value
T1	272.208074 μs
T2	326.176439 μs
Readout assignment error	0.003418
P(0 | 1)	0.005615
P(1 | 0)	0.001465
Readout length	2180.0 ns
Identity gate error	0.000157
Identity gate time	32.0 ns
R_z gate error	0.0
$\sqrt{X}$ (S_x) gate error	0.000157
$\sqrt{X}$ (S_x) gate time	32.0 ns
Pauli-X gate error	0.000157
Pauli-X gate time	32.0 ns
CZ gate error	0.001180
CZ gate time	68.0 ns

Table D1. Calibration metrics for the Heron R3 processor (`ibm_boston`)[104] recorded on 13 June 2026 at 13:18:06 IST. These values correspond to the device's pre-execution calibration snapshot, including coherence times (T_1, T_2), readout assignment errors, gate error rates, and gate durations. Median values are reported across all qubits and gate instances on the device. Please note that the quantum hardware is usually calibrated every 24 hours.

E Fragment orbital space decomposition for DMET-SQD simulations.

Molecule	Fragment	(o, e)	N_{orb}	$\lvert A_y\rvert$	$\lvert B_y\rvert$	$\lvert \text{Cor}_y\rvert$	$\lvert \text{Vir}_y\rvert$	$\lvert \text{Env}_y\rvert$
HCNO	[H]	(2, 2)	16	1	1	10	4	15
	[C]	(10, 10)		5	5	6	0	11
	[N]	(10, 10)		5	5	6	0	11
	[O]	(10, 10)		5	5	6	0	11
CH_3NO	[C]	(10, 10)	18	5	5	7	1	13
	[H]	(2, 2)		1	1	11	5	17
	[H]	(2, 2)		1	1	11	5	17
	[H]	(2, 2)		1	1	11	5	17
	[N]	(10, 10)		5	5	7	1	13
	[O]	(10, 10)		5	5	7	1	13
CH_5NO	[C]	(10, 10)	20	5	5	8	2	15
	[H]	(2, 2)		1	1	12	6	19
	[H]	(2, 2)		1	1	12	6	19
	[H]	(2, 2)		1	1	12	6	19
	[H]	(2, 2)		1	1	12	6	19
	[H]	(2, 2)		1	1	12	6	19
	[O]	(10, 10)		5	5	8	2	15
	[N]	(10, 10)		5	5	8	2	15
C_2H_3NO	[C]	(10, 10)	23	5	5	10	3	18
	[C]	(10, 10)		5	5	10	3	18
	[H]	(2, 2)		1	1	14	7	22
	[H]	(2, 2)		1	1	14	7	22
	[H]	(2, 2)		1	1	14	7	22
	[N]	(10, 10)		5	5	10	3	18
	[O]	(10, 10)		5	5	10	3	18
C_2H_5NO	[C]	(10, 10)	25	5	5	11	4	20
	[C]	(10, 10)		5	5	11	4	20
	[H]	(2, 2)		1	1	15	8	24
	[H]	(2, 2)		1	1	15	8	24
	[H]	(2, 2)		1	1	15	8	24
	[H]	(2, 2)		1	1	15	8	24
	[H]	(2, 2)		1	1	15	8	24
	[N]	(10, 10)		5	5	11	4	20
	[O]	(10, 10)		5	5	11	4	20
CH_4N_2O	[C]	(10, 10)	24	5	5	11	3	19
	[N]	(10, 10)		5	5	11	3	19
	[H]	(2, 2)		1	1	15	7	23
	[H]	(2, 2)		1	1	15	7	23
	[H]	(2, 2)		1	1	15	7	23
	[H]	(2, 2)		1	1	15	7	23
	[N]	(10, 10)		5	5	11	3	19
	[O]	(10, 10)		5	5	11	3	19
NOCl	[N]	(8, 10)	19	5	3	11	0	14
	[O]	(8, 10)		5	3	11	0	14
	[Cl]	(12, 18)		9	3	7	0	10
HOSCN	[H]	(2, 2)	25	1	1	18	5	24
	[O]	(10, 10)		5	5	14	1	20
	[S]	(15, 18)		9	6	10	0	16
	[C]	(10, 10)		5	5	14	1	20
	[N]	(10, 10)		5	5	14	1	20

Table E1. For each fragment: impurity active space (o, e) (spatial orbitals and electrons); $\lvert A_y\rvert$, fragment orbitals; $\lvert B_y\rvert$, bath orbitals ($\lvert B_y\rvert \leq \lvert A_y\rvert$); $\lvert \text{Cor}_y\rvert$, fully occupied environment orbitals; $\lvert \text{Vir}_y\rvert$, empty environment orbitals; $\lvert \text{Env}_y\rvert = \lvert B_y\rvert + \lvert \text{Cor}_y\rvert + \lvert \text{Vir}_y\rvert$, total environment orbitals at $\varepsilon_{occ} = 10^{-13}$.

F Classical Resources for HOSCN molecule for varying ε_{occ}

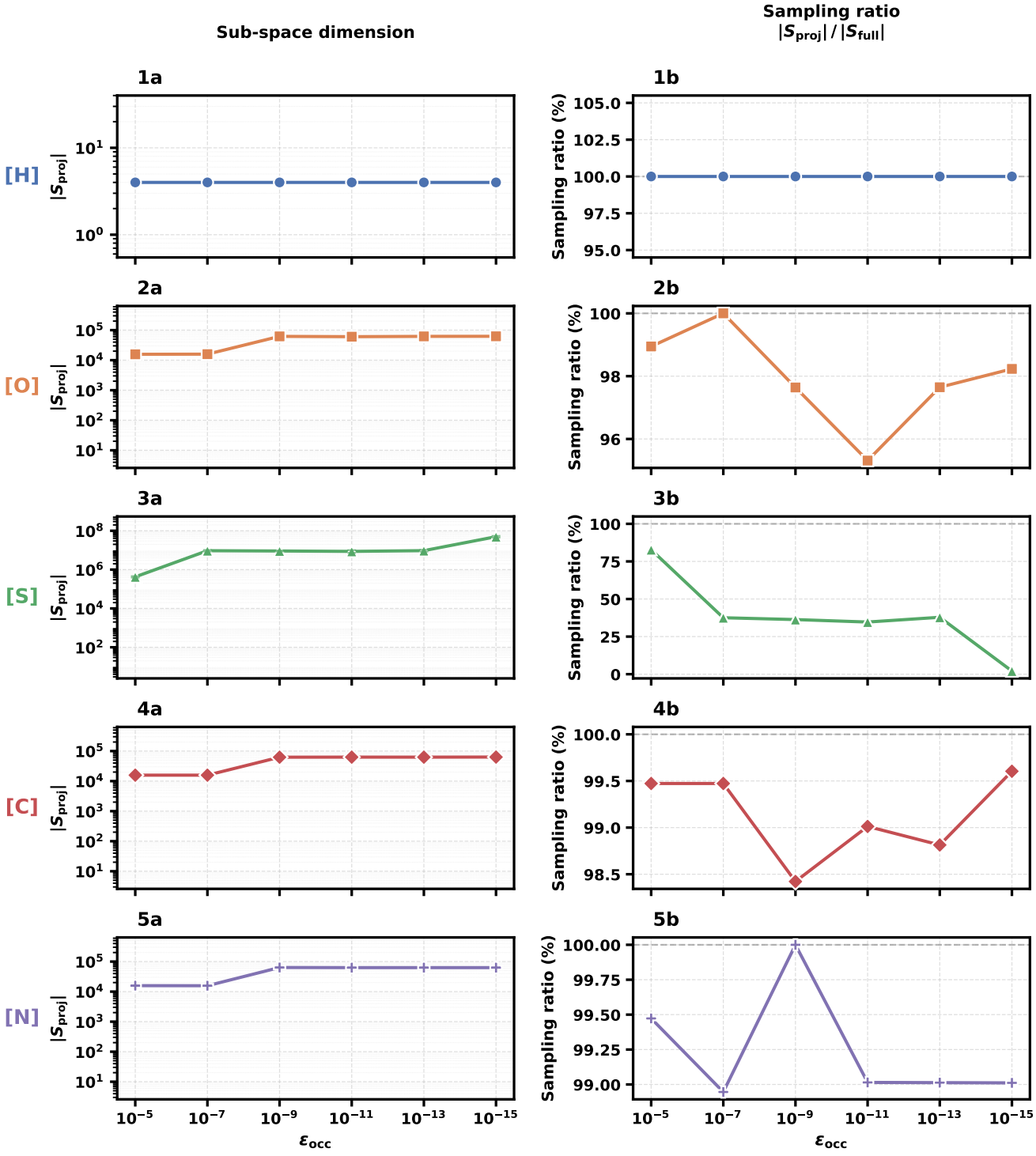

Figure F1. Classical S-CoRe resources for HOSCN/STO-3G as a function of ε_{occ} : Sub-space dimension $|S_{\text{proj}}|$ (a), and sampling ratio $|S_{\text{proj}}|/|S_{\text{full}}|$ (b) for each DMET fragment of HOSCN across different ε_{occ}, executed on IBM Heron R3 Boston.